Geochemistry, Geophysics, Geosystems®

AGU ADVANCING EARTH AND SPACE SCIENCES

METHOD
10.1029/2023GC011234

Key Points:

- We give a suite of MATLAB scripts (LIME) that implements an Aitchison (1982, https://doi.org/10.1111/j.2517-6161.1982.tb01195.x) log-ratio approach for mass-balance calculations in petrology
- Our algorithm constrains phase proportions to positive values and provides errors that consider known uncertainties of input compositions
- We demonstrate the utility of propagated phase uncertainties through calculations of magma crystallinity and melting reaction coefficients

Correspondence to:
K. Prissel,
kelsey.prissel@nasa.gov

Citation:
Prissel, K., Olive, J.-A., & Krawczynski, M. J. (2023). A log-ratio-based algorithm for petrologic mass-balance problems and uncertainty assessment. *Geochemistry, Geophysics, Geosystems*, *24*, e2023GC011234. https://doi.org/10.1029/2023GC011234

Received 12 SEP 2023
Accepted 25 NOV 2023

A Log-Ratio-Based Algorithm for Petrologic Mass-Balance Problems and Uncertainty Assessment

Kelsey Prissel[1,2,3], Jean-Arthur Olive[4], and Michael J. Krawczynski[3]

[1]Jacobs, NASA Johnson Space Center, Houston, TX, USA, [2]Earth and Planets Laboratory, Carnegie Institution for Science, Washington, DC, USA, [3]Department of Earth and Planetary Sciences, Washington University in St. Louis, St. Louis, MO, USA, [4]Laboratoire de Géologie, CNRS—École normale supérieure—PSL University, Paris, France

Abstract

We provide a new algorithm for mass-balance calculations in petrology and geochemistry based on the log-ratio approach championed initially by John Aitchison (e.g., Aitchison, 1982, https://doi.org/10.1111/j.2517-6161.1982.tb01195.x; Aitchison, 1984, https://doi.org/10.1007/bf01029316) along with the underlying principles, mathematical frameworks, and data requirements. Log-ratio Inversion of Mixed End-members (LIME) is written in MATLAB and calculates phase proportions in an experiment or rock given a bulk composition, the composition of each phase, and the associated compositional uncertainties. An important advantage of LIME is that performing the mass-balance calculation in inverse log-ratio space constrains phase proportions to be between 0 and 100 wt.%. Further, the resulting LIME phase proportions provide realistic estimates of uncertainty regardless of data distribution. These two characteristics of LIME improve upon standard multiple linear regression techniques, which may yield negative values for phase proportions if non-constrained or report oversimplified symmetric errors. Primary applications of LIME include estimating phase abundances, calculating melting and metamorphic reaction stoichiometries, and checking for open system behavior in phase equilibria experiments. The technique presented here covers whole-rock analysis, mineralogy, and phase abundance, but could be extended to isotopic tracers, trace element modeling, and regolith component un-mixing. We highlight the importance of uncertainty estimations for phase abundances to the fields of petrology and geochemistry by comparing our results from LIME to previously published mass-balance calculations. Furthermore, we present case studies that demonstrate the role of mass-balance calculations in determining magma crystallinity and defining melting reactions.

Plain Language Summary

"The principle of mass balance is simply: some of it, plus the rest of it, equals all of it."—Stormer, Jr. and Nicholls (1978, https://doi.org/10.1016/0098-3004(78)90083-3). In the study of igneous and metamorphic rocks, mineral abundances are used to classify rock types, characterize volcanic processes, and define chemical reactions. Common approaches for calculating phase abundances, such as multiple linear regression, suffer from mathematical artifacts that can lead to inaccurate assumptions (e.g., Chayes, 1960, https://doi.org/10.1029/jz065i012p04185; Miesch, 1969, https://doi.org/10.1007/978-1-4615-8633-3_10). Compositions are reported as percentages with sums mathematically constrained to 100%. This fixed sum constraint may lead to erroneous coefficients when determining chemical reactions (e.g., a mineral being on the wrong side of the chemical reaction). Additionally, mineral abundances are often reported without any estimate of uncertainty. This is problematic when those values are used in subsequent calculations to define trends without statistical significance. We have developed a suite of MATLAB scripts to calculate phase abundances in natural rocks and experimental samples. This algorithm, referred to as LIME, improves upon previous methods by constraining phase proportions to positive values and providing realistic estimates for uncertainty on each phase abundance. We demonstrate the utility of these improvements in two case studies related to magma crystallization and mantle melting.

1. Introduction

Mass-balance calculations play a crucial role in petrology and geochemistry, serving as a powerful tool for understanding the distribution, transformation, and cycling of elements in petrologic systems, as well as assessing possible open-system behavior in experimental phase equilibria results. By quantifying elemental budgets, mass-balance calculations enable us to decipher the intricate chemical dynamics that shape the composition of Earth and planetary materials. Mass-balance calculations quantify the fluxes of elements between reservoirs and elucidate the factors controlling elemental distributions in a system. The proportions of phases, even if

macroscopically out of equilibrium, constrain processes occurring in diverse geological settings, such as magmatic systems or metamorphic regimes.

In petrology, mass-balance calculations provide insight into the genesis and evolution of rocks and minerals. By examining the bulk chemical composition of rocks and comparing them with the expected compositions based on mineral equilibria, thermodynamics, and known phase relations, mass-balance calculations—when correctly done—discern the processes that have influenced the rock's formation and subsequent modifications. Petrological mass-balance studies have shed light on magma differentiation, viscosity, fractional crystallization, assimilation, magma mixing, and metasomatic processes among other processes (e.g., DePaolo, 1981; Langmuir et al., 1992; Marsh, 1981).

User-friendly implementations for mass-balance calculations have served the petrology community for decades (e.g., XLFRAC, Stormer, Jr. and Nicholls (1978); Simplex, Baker et al. (1994); GeoBalance, Li et al. (2020)). Bryan et al. (1969) first applied the least-squares approximation to the issue of mass balance in petrology. The standard regression technique has been improved by approaches that incorporate analytical uncertainty and/or constrain solutions to non-negative values (e.g., Albarede & Provost, 1977; Reid et al., 1973; Wright & Doherty, 1970). With geochemical data from igneous and metamorphic rocks, the mass-balance problem is often over-determined (i.e., more oxide constraints than phases), and thus methods have also been developed to mitigate this issue for least-squares mass-balance calculations (e.g., Ghiorso, 1983; Stormer, Jr. & Nicholls, 1978). The least-squares approach yields phase proportion uncertainties that are determined based on the residual of the regression fit. More recently, Monte Carlo error propagation for mass-balance calculations has been implemented by the GeoBalance Excel VBA program (Li et al., 2020) and Python-based MassBalanceCal (Zhang et al., 2023). Though additional mathematical improvements for compositional mass-balance exist (e.g., Aitchison, 1982; Greenacre et al., 2023), they have yet to be specifically tailored for practical application in petrology.

Here, we present a new algorithm (Log-ratio Inversion of Mixed End-members, or LIME) for mass-balance calculations in petrology and geochemistry based on the log-ratio approach initially championed by John Aitchison (e.g., Aitchison, 1982). This approach to mass-balance constrains solutions to non-negative values, incorporates analytical uncertainty, and yields asymmetric error bars that realistically reflect compositional uncertainties in the bulk and phase compositions. We demonstrate improvements to mass-balance calculations through case study applications, and highlight the significance of phase abundances in elucidating petrologic processes. Primary applications of LIME include determining phase abundances, calculating melting and metamorphic reaction stoichiometries, and checking for closed/open system behavior in phase equilibria experiments. The technique presented here covers whole-rock analysis, mineralogy, and phase abundance, but could be extended to isotopic tracers, trace element modeling, or regolith component un-mixing (e.g., Korotev & Kremser, 1992).

2. A Log-Ratio Approach to the Mass-Balance Problem

2.1. Mathematical Formulation and Challenges of the Mass-Balance Problem

Consider a rock made up of P mineral phases, each being analyzed for C element concentrations (i.e., major oxides, in wt.%). The bulk composition of the rock is also being determined, and written as a C-element column vector b. Assuming homogeneous phases and perfect measurements, we have the following linear relation

$$b = Ep \tag{1}$$

where E is the $C \times P$ matrix whose columns represent the major oxide composition of each mineral phase (E_{ij} denotes the proportion of oxide i in the jth mineral phase), and p is the P-element column vector containing the relative proportions of the mineral phases. Equation 1 describes the forward problem we wish to invert while accounting for measurement uncertainty and phase inhomogeneity. Knowing the bulk composition b and the end-member composition E with some prior uncertainty, we wish to identify the values of vector p that best predict b, as well as the associated uncertainty on p.

An intuitive way to do so is to perform a least-squares regression on matrix E and data vector b in order to estimate which vector $p = p^*$ minimizes the distance between Ep and b (in the L^2 norm sense: $\|Ep - b\|_2$) (Bryan et al., 1969; Wright & Doherty, 1970). Such methods rely on the assumption that uncertainties on b and p follow a Gaussian distribution. Complications arise, however, when one attempts to perform least-squares inversion on

compositional vectors and matrices that carry only relative information (proportions) such as b, p, and E. Indeed, such vectors are subject to a positivity and unit-sum constraint:

$$b_i > 0; \sum_{i=1}^{C} b_i = 1 \tag{2}$$

$$p_i > 0; \sum_{i=1}^{P} p_i = 1 \tag{3}$$

$$E_{ij} > 0; \sum_{i=1}^{C} E_{ij} = 1 \tag{4}$$

Because of these constraints, the L^2 norm does not constitute an appropriate measure of distance in a simplex, that is, a space of compositional vectors. This means that compositional uncertainties, which are related to the spread of compositional distributions, are not properly assessed when one uses usual metrics such as the standard deviation. Besides, a C-simplex (S^C) spanned by compositional vectors with C components really has $C-1$ degrees of freedom (i.e., once the $C-1$ first components have been assigned, the unit-sum constraint sets the value of the Cth-component). Multivariate Gaussian distributions can therefore not be defined on a C-simplex because their natural domain of definition is the Euclidian space $\mathbb{R}^{C-1}$. By ignoring these facts when performing linear regressions on compositional data sets, researchers often end up with skewed results, such as negative phase contents and misleading (or absent) measures of uncertainty (e.g., Pawlowsky-Glahn et al., 2015).

Albarede and Provost (1977) recognized these limitations and proposed to perform standard least-squares methods not on compositions directly, but on a transformed data set that would not be subject to the constraints listed above. Specifically, they worked with vectors $\Omega_i \in \mathbb{R}^C$ instead of compositional vectors $\omega_i \in S^C$, where $\sin^2\Omega_i = \omega_i$ (Ray & Szekely, 1973). While this approach allows the definition of Gaussian uncertainties on the transformed vectors, it does not get rid of spurious correlations inherited from the unit-sum constraint.

A more general approach was put forward by Aitchison (1982), who established the field of compositional geometry. He defined a set of basic operations on the simplex S^C providing it a structure analogous to that of Euclidian spaces such as $\mathbb{R}^{C-1}$. Specifically, through the definition of an inner-product in S^C, he introduced a measure of distance in the simplex now referred to as Aitchison distance. He also introduced one-to-one mappings between S^C and $\mathbb{R}^{C-1}$. The first one, the additive log-ratio transform (alr), consists of choosing a component of a compositional vector, and taking the logarithm of each of the other components divided by the chosen component. This yields a vector of real numbers that is subject to neither the positivity nor unit-sum constraints, making it suitable for standard regression techniques, as well as other statistical tests that require unconstrained real space. One drawback of this transform is that is does not provide a direct correspondence between the Aitchison distance and the Euclidian distance. In other words, taking the Aitchison norm of compositional vector x will not yield the same result as taking the L^2 norm of alr(x), which makes the interpretation of transformed data less straightforward (Egozcue et al., 2003). Aitchison (1982) also introduced the centered-log-ratio (clr) transform, which consists of the logarithm of each component of a composition vector divided by the geometric mean of all its components (see Appendix A). This transform allows a direct correspondence between Aitchison and Euclidian distances. However, the clr of a C-component composition yields a C-component vector subject to a zero-sum constraint. This implies that clr-transformed data still carry spurious correlations. Specifically, any covariance matrix defined on clr-transformed data sets will be singular, which is a major drawback for many least-squares algorithms.

A third transform, the isometric log-ratio transform (ilr) was introduced by Egozcue et al. (2003). It is based on the definition of an orthogonal basis in the simplex in the Aitchison sense (see Appendix A). The ilr allows a direct correspondence (isometry) between Aitchison and Euclidian distances, meaning that the L^2 norm of ilr(x) yields the same result as the Aitchison norm of the C-element compositional vector x. In addition, the components of ilr(x) are not subject to any sum or positivity constraint, and span an unconstrained, Euclidian space $\mathbb{R}^{C-1}$. These properties make the ilr-transform a powerful tool for compositional data analysis. A now standard practice consists of taking the ilr of compositional data sets and performing standard statistical or inverse methods to the transformed data sets (Chave, 2017). Here we apply this methodology to inverting the convex linear mixing problem (Equation 1) and present a suite of numerical programs tailored for petrological mass-balance calculations.

2.2. The Log-Ratio Transformed Mass Balance Problem

The very first step in our methodology is to eliminate any zero present in our data sets. We do so by following the protocol detailed in Appendix B. We then reformulate Equation 1 as a non-linear problem involving the isometric log-ratio (ilr) transform of each compositional vector involved:

$$\hat{b} = \Psi\left(\hat{E}, \hat{p}\right). \tag{5}$$

The hat notation denotes the ilr-transform. We define the ilr of a matrix as the matrix composed of the ilr of each of its columns. $\hat{p}$ is a vector with $P-1$ real components not subject to the unit sum constraint, and $\hat{E}$ is a $(C-1)\times P$ matrix. Ψ denotes a non-linear operator which, from Equation 1, explicitly writes

$$\Psi\left(\hat{E}, \hat{p}\right) = ilr^{-1}\left(\hat{E}\right) ilr^{-1}(\hat{p}). \tag{6}$$

A convenient way of handling uncertainties on end-member compositions is to treat $\hat{E}$ as a model parameter (like $\hat{p}$) rather than as data (like $\hat{b}$). We, therefore, assemble a column vector $\hat{m}$ with $\mathrm{CP}-1$ components that contains all the columns of $\hat{E}$ taken sequentially, followed by the $P-1$ elements of $\hat{p}$. Following a standard Bayesian approach, we treat our initial estimate of $\hat{m}$ (model vector) as a random drawing from a multivariate Gaussian distribution f_{PRIOR}. This prior distribution is centered on an a-priori estimate of $\hat{m}$ termed $\hat{m}_0$, which contains the actual measurements of end-member compositions (columns of $\hat{E}$) and an initial "guess" of $\hat{p}$ which assumes all mineral phases are in equal proportions. The spread of f_{PRIOR} is expressed through a $(\mathrm{CP}-1)\times(\mathrm{CP}-1)$ covariance matrix C_{PRIOR} containing (a) uncertainty related to the measurements of the end-member compositions ($P(C-1)\times P(C-1)$ upper-left block), and (b) uncertainty on the initial guess of the phase proportions ($(P-1)\times(P-1)$ lower-right block), taken to be large (see Appendix C):

$$f_{PRIOR}(\hat{m}) \propto \exp\left(-\frac{1}{2}(\hat{m}-\hat{m}_0)^T C_{PRIOR}(\hat{m}-\hat{m}_0)\right). \tag{7}$$

Similarly, we model measurement uncertainty on the ilr of the bulk composition vector $\hat{b}$ as a Gaussian distribution g_{OBS} centered on the measured ilr-bulk composition $\hat{b}_0$, with a $(C-1)\times(C-1)$ covariance matrix C_D (see Appendix C for details on constructing C_D).

$$g_{OBS}\left(\hat{b}\right) \propto \exp\left(-\frac{1}{2}\left(\hat{b}-\hat{b}_0\right)^T C_D\left(\hat{b}-\hat{b}_0\right)\right) \tag{8}$$

A slight rewriting of Ψ in Equation 5 allows us to define operator Φ so that

$$\hat{b} = \Phi(\hat{m}) \tag{9}$$

Equation 9 is a concise form of the forward mass-balance problem which has been rendered non-linear by the use of the ilr-transform.

2.3. Inversion Procedure

We apply a quasi-Newton iterative method (Tarantola, 1981) to invert the non-linear problem described by Equation 9. This method starts from the prior distribution $f_{PRIOR}(\hat{m})$ (characterized by covariance matrix C_{PRIOR} and centered on $\hat{m}_0$) and employs discrete steps to approach a posterior distribution $f_{POST}(\hat{m})$ that allows the best prediction of the data distribution $g_{OBS}\left(\hat{b}\right)$ in the L_2-norm sense. In practice, we iterate over the model vector $\hat{m}^{(n)}$ after initializing it at $\hat{m}^{(0)} = \hat{m}_0$. At each iteration n, we estimate the Jacobian matrix of Φ through a centered finite-difference approximation

$$J_{kl}^{(n)} = \frac{\partial \hat{b}_k}{\partial \hat{m}_l} = \frac{\Phi_k(\hat{m}_l + \Delta\hat{m}/2) - \Phi_k(\hat{m}_l - \Delta\hat{m}/2)}{\Delta\hat{m}} + O\left(\Delta\hat{m}^2\right) \tag{10}$$

Discrete steps of $\Delta\hat{m} = 10^{-3}$ were found sufficiently small to allow a robust estimate of J in most petrological applications. The Jacobian is then used in the following steepest-descent scheme:

$$\hat{m}^{(n+1)} = \hat{m}^{(n)} - \mu K^{(n)}\left[J^{(n)T} C_D^{-1}\left(\Phi\left(\hat{m}^{(n)}\right) - \hat{b}\right) + C_{PRIOR}^{-1}\left(\hat{m}^{(n)} - \hat{m}_0\right)\right] \tag{11}$$

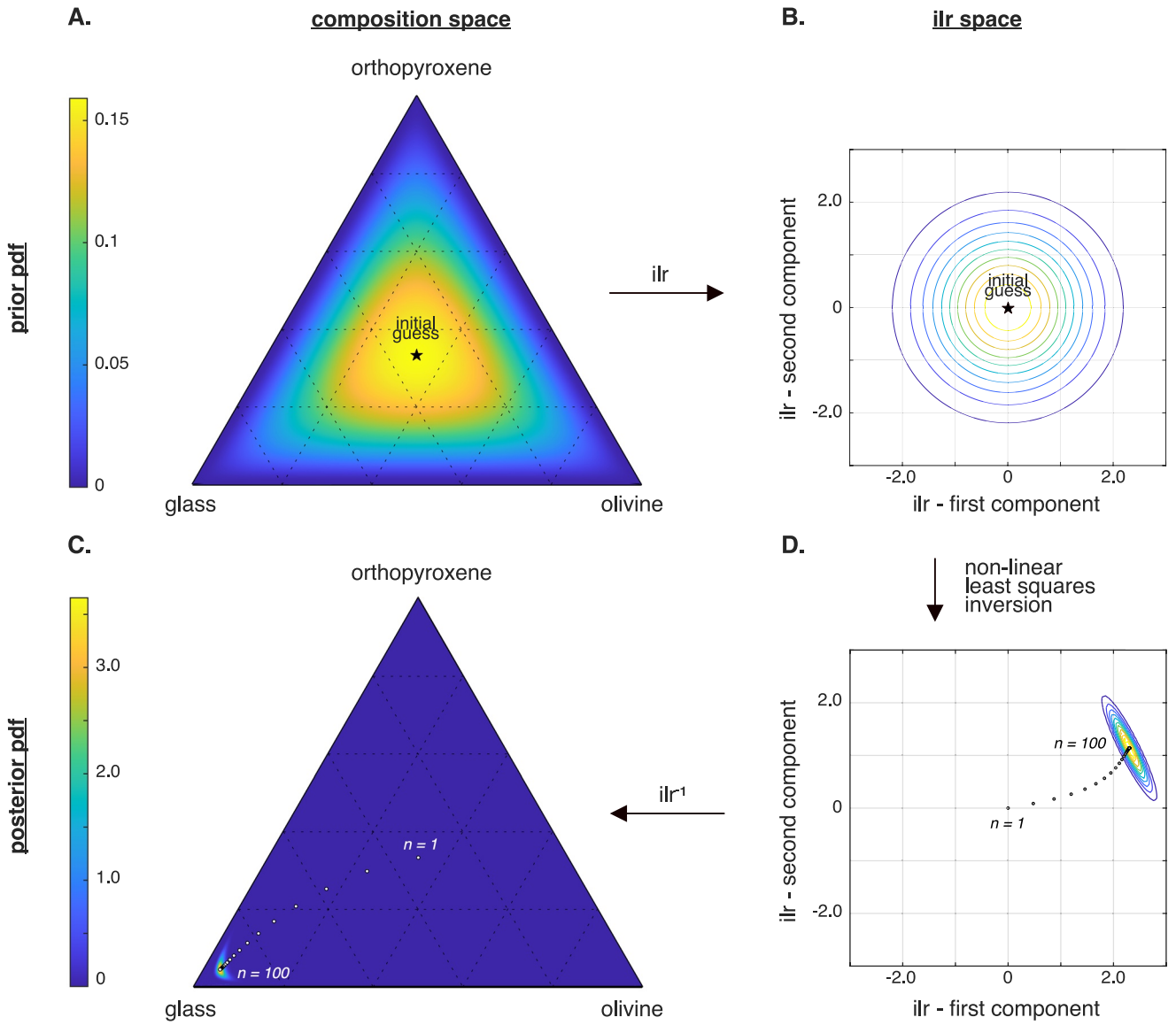

Figure 1. Illustration of the LIME algorithm in composition space (left column) and ilr-transformed space (right column), based on experiment C321 from Krawczynski and Grove (2012) (see Section 3). LIME builds a probability density function (PDF) on the mixture (here, proportions of glass, olivine, and plagioclase) that best explains a known bulk composition, accounting for uncertainties on both the bulk and end-member compositions. The algorithm initially postulates equal phase proportions, with large model uncertainties (a). In ilr-space, this corresponds to a wide, isotropic Gaussian distribution centered on zero (b). An iterative procedure is then applied in ilr-space to find the best-fitting phase proportions and associated probability density function (d). It typically converges within tens of iterations (circles show progression from iteration $n = 1$ to $n = 100$). The results are then converted back to composition space (c). The same results are plotted as PDFs of individual phase proportions in Figure 2a.

where

$$K^{(n)} = \left[J^{(n)T} C_D^{-1} J^{(n)} + C_{PRIOR}^{-1}\right]^{-1} \tag{12}$$

and μ is a damping term we set equal to 0.9 by default. We empirically find that a smaller value sometimes improves convergence. $\hat{m}^{(n)}$ typically converges within tens of iterations (e.g., Figure 1) toward a model vector $\hat{m}^{(\infty)}$, which contains our best estimate of the phase proportions.

By default, our method only performs the updates described by Equation 11 on the components of the model vector that describe phase proportions, that is, we do not update the values that describe the composition of the end-member phases. Doing so assumes the end-member compositions and their uncertainties are well described

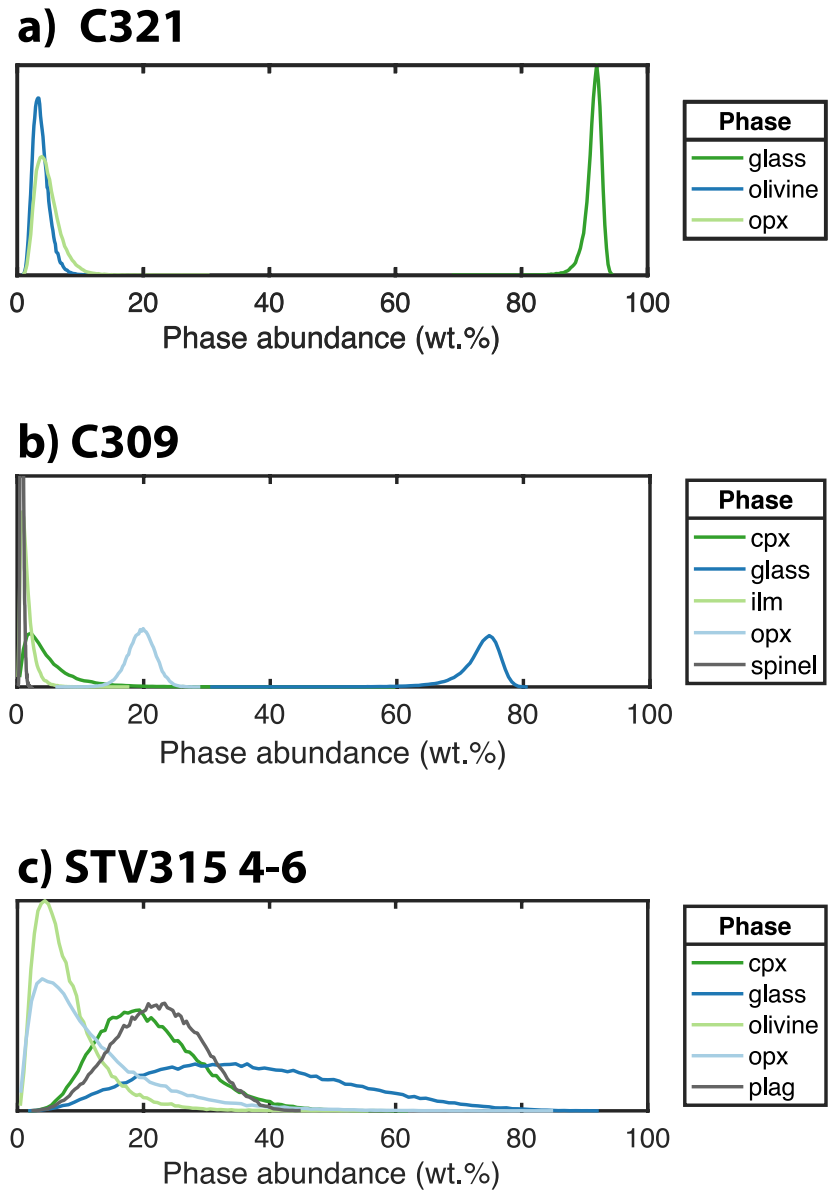

Figure 2. LIME output probability density functions for computed phase abundance results on three experiments: C321 (a) and C309 (b) from Krawczynski and Grove (2012), and STV315 4–6 (c) from Pichavant and Macdonald (2007). For numerical LIME results and comparison to published phase abundances, see Table 1.

by the input data set, and limits the number of parameters to invert for. To explicitly treat the end-member compositions as model parameters, users can switch on the option `yn_iterate_on_endmembers`.

The uncertainty associated with our best model estimate is described by a Gaussian distribution f_{POST} centered on $\hat{m}^{(\infty)}$ with covariance matrix C_{POST}:

$$f_{POST}(\hat{m}) \propto \exp\left(-\frac{1}{2}\left(\hat{m}-\hat{m}^{(\infty)}\right)^{T} C_{POST}\left(\hat{m}-\hat{m}^{(\infty)}\right)\right), \tag{13}$$

where

$$C_{POST} = C_{PRIOR} - C_{PRIOR}J^{(\infty)T}\left(J^{(\infty)}C_{PRIOR}J^{(\infty)T} + C_D\right)^{-1}J^{(\infty)}C_{PRIOR}. \tag{14}$$

Table 1
Comparison of LIME Results and Published Values From Krawczynski and Grove (2012) Experiments C321 and C309, Pichavant and Macdonald (2007) Experiment STV315 4-6, and Dawson and Krawczynski (2022) Experiment L007

		LIME	
Phase	Reported (wt.%)	Mean ± s.d. (wt.%)	Best fit (wt.%)
Krawczynski and Grove (2012) C321			
Glass	93	91 ± 1	$92.0\pm^{0.3}_{1.1}$
Orthopyroxene	4	5 ± 2	$4\pm^{1}_{1}$
Olivine	2	4 ± 1	$3.6\pm^{0.9}_{0.7}$
Krawczynski and Grove (2012) C309			
Glass	100	73 ± 3	$74\pm^{1}_{3}$
Orthopyroxene	27	20 ± 2	$20\pm^{1}_{2}$
Clinopyroxene	−22	5 ± 4	$4\pm^{3}_{2}$
Ilmenite	−7	2 ± 1	$1.3\pm^{0.7}_{0.5}$
Spinel	1	0.8 ± 0.2	$0.7\pm^{0.2}_{0.1}$
Pichavant and Macdonald (2007) STV315 4-6			
Glass	64	37 ± 16	$38\pm^{10}_{12}$
Plagioclase	12	22 ± 7	$24\pm^{3}_{6}$
Clinopyroxene	22	21 ± 8	$22\pm^{4}_{7}$
Orthopyroxene	−10	12 ± 10	$10\pm^{6}_{5}$
Olivine	12	8 ± 5	$7\pm^{3}_{3}$
Dawson and Krawczynski (2022) L007			
Glass		42 ± 6	$43\pm^{3}_{5}$
Amphibole		25 ± 4	$25\pm^{2}_{3}$
Orthopyroxene		16 ± 2	$16\pm^{1}_{2}$
Clinopyroxene		10 ± 4	$10\pm^{3}_{2}$
Plagioclase		7 ± 6	$5\pm^{4}_{2}$

Note. Here, "Best Fit" gives the best fitting phase abundance (peak center) and phase abundance increments to the 25th (−) and 75th (+) percentiles of the posterior distribution. Users can define which percentiles of interest to calculate before running LIME. For a Gaussian distribution, percentiles of 15.9 and 84.1 correspond approximately to 1σ.

We sample this distribution in composition space by drawing 80,000 random samples from f_{POST} and then computing their inverse-ilr transform. This allows us to construct distribution curves for each mineral phase that are bounded between 0 and 1 (Figure 2). The number of random samples from f_{POST} can be varied in the code, but we found 80,000 was enough to generate smooth probability density curves for phases in most situations. This value can be increased or decreased by the user (see Appendix E).

The above methodology is implemented as a MATLAB (The MathWorks Inc, 2022) code openly distributed on the Zenodo repository (Prissel et al., 2023), and described in Appendix E. User-provided data sets of end-member compositions and bulk compositions are first processed to eliminate any zero not compatible with the calculation of the ilr transform (Appendix A). This is achieved with the method highlighted in Appendix B. Data set and model uncertainties (expressed in the form of covariance matrices C_D and C_{PRIOR}) are assessed as detailed in Appendix C. Finally, Appendix D summarizes how synthetic data sets were used to evaluate the code's performance.

3. Examples and Advantages of Using LIME for Mass-Balance Calculations

3.1. Constrained Phase Proportions From 0% to 100%

To demonstrate how the mass-balance calculations of LIME compare to previous methods, we detail three examples in the following text, Table 1 and Figure 2. As described in Section 2, an important advantage of LIME is that performing the mass balance calculation in ilr-space constrains phase proportions to be between 0 and 100 wt.%. Further, the resulting probability density curves give phase proportions as well as realistic, and often asymmetric, uncertainties on the proportion for each phase. These two characteristics of LIME improve upon the commonly used multiple linear regression techniques, which only report symmetric errors and may give standard deviations that yield nonsensical negative values for phase proportions. Owing to these improvements, several authors have implemented previous versions of LIME (Krawczynski & Olive, 2011) in calculating the phase proportions of igneous and metamorphic rocks (Codillo et al., 2022; Collinet et al., 2015; Davis & Hirschmann, 2013; Gavrilenko et al., 2019; Grove et al., 2013; Guenther et al., 2022; Mitchell & Grove, 2015; Stadermann et al., 2022). LIME has also been used to assess the open system behavior of iron in experimental samples (Brown & Grove, 2015; Goltz et al., 2022; Prissel et al., 2018). We describe how to perform the Fe loss calculation in Appendix E5.

Inputs required for a LIME calculation are the bulk composition of the experiment and individual analyses of each experimental phase. LIME will transform each individual analysis into ilr-space and then compute a geometric mean, so it is important to input multiple individual analyses instead of an average composition. We provide user instructions and a detailed discussion of input data and uncertainty in Appendix E. We have also provided the compositional data input files for example, experiments C321, C309, L007, and STV315 4–6 in the LIME source files (https://github.com/kprissel/LIME).

The main output of the LIME calculation is a probability density function of phase proportions (Figure 2). We present published experiment C321 from Krawczynski and Grove (2012) as a simple example of the LIME output. This experiment contains three phases: olivine, orthopyroxene, and glass (Figures 1 and 2a). The mass balance for this experiment is relatively straightforward because all of the phases are homogeneous and there were no issues with open system behavior in the experiment (i.e., Fe loss). The area under each curve is normalized to 1. For experiment C321, the peaks are relatively symmetric, with approximate normal distributions and low

uncertainties. For each phase, the peak center and width relate to the best fitting phase proportion and the uncertainty on that value, respectively. The original publication calculated the mass balance with a non-constrained linear least squares method, and there is good agreement between the reported phase proportions and those calculated with LIME (Table 1).

For petrologic mass-balance calculations it has been common practice to either report negative coefficients (e.g., Krawczynski & Grove, 2012; Pichavant & Macdonald, 2007) or not report mass-balance coefficients when the results are negative (e.g., Canil & Bellis, 2008). To demonstrate the improvement over common regression methods used for the estimation of phase proportions, we provide two examples of experiments with previously reported negative phase abundances (Figure 2, Table 1).

Experiment C309 from Krawczynski and Grove (2012) has published negative values for mass balance coefficients of clinopyroxene and ilmenite phases. Standard linear regression techniques are not constrained to give positive phase proportion coefficients, resulting in non-intuitive negative abundances for existing phases. The LIME algorithm constrains the mass balance calculation to positive space, giving best-fit proportions that more realistically reflect the observed phase assemblage. The phase proportion coefficients for C309 are reported in Table 1 along with the first and third quartiles from the probability density function, which emphasize the asymmetric nature of the probability functions when mapped back into non-Euclidian composition space (Figure 2b). Similarly, results for a recalculation of experiment STV315 4-6 from Pichavant and Macdonald (2007) demonstrate the difference between linear regression methods and the LIME algorithm (Figure 2c). Specifically, for this experiment, there is a drastic difference in the calculated "glass" percents, which would have significant impact when using the glass proportion to define crystallization trends or determine the total crystallinity of an experiment.

3.2. Best Practices for Reporting Uncertainty From LIME

The most accurate representation of uncertainty in LIME derives from the posterior distribution of the randomly sampled points in ilr-space. The uncertainty on each phase proportion is graphically depicted as a standard LIME output (Figure 2). We recognize that it might be impractical to report uncertainty as a figure, especially in papers with dozens of these calculations and bulk compositions with many phases. In addition, phase abundances are often used as inputs into subsequent calculations (e.g., Section 4) and propagating uncertainty would be impossible from a figure. Therefore, we have developed multiple ways to numerically express uncertainty when using LIME, and we include two examples in Table 1.

The output of LIME gives three estimations for uncertainty on phase proportions. The simplest estimate of uncertainty is the average and 1σ standard deviation of the 80,000 random samples computed in the LIME algorithm (i.e., "Mean $\pm$ s.d." in Table 1). Because the standard deviation will only provide symmetric error, these values are an oversimplification and will be incorrect if the distribution is not normal, hindering any predictive power for this measure of error. For the cases where the probability density functions approach normal distributions, the LIME mean and standard deviation are similar to the previously reported uncertainties. However, the mean and standard deviation should not be used for cases that have significantly skewed distributions.

If using phase proportion data in subsequent calculations, the highest fidelity uncertainty propagation would use the 80,000 random sample values in a Monte Carlo method. One output of the LIME algorithm is a vector R that contains all 80,000 of the randomly sampled points in ilr-space converted back to common composition space. This vector can be used in subsequent calculations, such as total crystallinity during a series of fractional crystallization experiments. This is the suggested method to achieve the full benefit of our methodology, and R is given as the second tab, "Modeled abundances," in the output spreadsheet (see Appendix E). An important consideration is that, because these samples are randomly generated, the resulting R values will not be exactly the same in each calculation. However, through numerous tests we have determined that the resulting phase proportions and uncertainties are reproducible to 0.1 wt.% when re-running the LIME algorithm for identical inputs.

For numerically reporting uncertainties from LIME, we recommend reporting the best fitting composition (peak) and boundaries for user-defined percentiles of the probability density functions. This method incorporates the asymmetric nature of the phase proportion curves and reports an upper and lower bound instead of a single 1σ value (Table 1). For the recommended percentile approach, the user defines what percentiles will be used (here, 25th and 75th) to calculate the error (Appendix E). Uncertainty can be recalculated for different percentiles

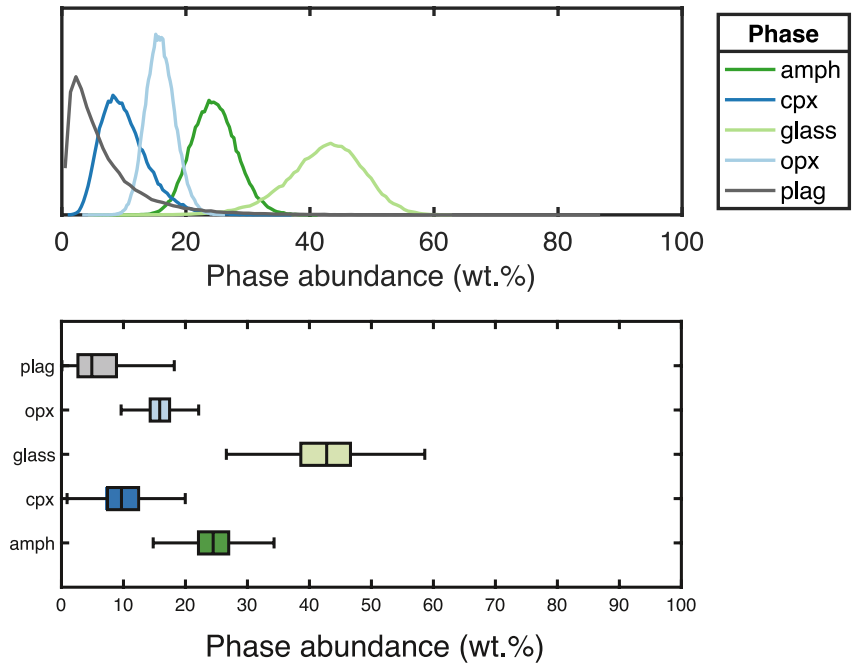

Figure 3. LIME output probability density function for computed phase abundance results on experiment L007 from Dawson and Krawczynski (2022). The results are also depicted as a box and whisker plot showing the best fitting phase abundances (peak center) and the 25th and 75th percentiles (Table 1).

by either changing the input percentiles and re-running the code, or by calculating the distribution from the random values in output *R*. As an example, numerically reported asymmetric uncertainties are compared to the probability density function for experiment L007 from Dawson and Krawczynski (2022) in Figure 3. All three methods of reporting uncertainty are output upon each LIME calculation.

4. Application to Petrologic Problems

4.1. Case Study 1: Magma Crystallinity

The crystallinity of a magma affects all of its physical properties including but not limited to density, viscosity, and volatile solubility. Therefore, studies of the dynamics and eruptibility of magma often rely on quantitative estimates for the total crystallinity of a magma (e.g., Blundy & Cashman, 2001; Blundy et al., 2008; Cashman, 1992; Cashman & Marsh, 1988; Cashman et al., 2017; Huber et al., 2012; Parmigiani et al., 2014). To demonstrate how LIME can be used to address questions pertaining to magma crystallinity, we have calculated the modal abundance of phases in a suite of amphibole-bearing experiments and have used the resulting proportions and uncertainties to quantify the change in crystallinity as a function of temperature.

A drastic increase in crystallinity due to the onset of amphibole crystallization has been hypothesized as the cause for hydrous basaltic magmas becoming trapped near the base of the Earth's crust rather than ascending to the surface (e.g., Barclay & Carmichael, 2004). Numerous experimental studies have demonstrated that the proportion of melt in an experiment exhibits an abrupt decrease (>30 wt.%) over a short cooling interval (20–50°C) once amphibole becomes a stable phase in the crystallizing assemblage (Barclay & Carmichael, 2004; Blatter et al., 2013; Foden & Green, 1992; Grove et al., 1997; Helz, 1973; Holloway & Burnham, 1972; Moore & Carmichael, 1998; Nandedkar et al., 2014; Sisson & Grove, 1993). This increase in amphibole abundance over a finite temperature interval, accompanied by a decrease in olivine and pyroxene abundance, reflects the peritectic reaction occurring between melt and clinopyroxene ± olivine to produce amphibole.

Variation in melt fraction (F) with change in temperature (T), or dF/dT, is a measure with which to quantify the change in crystallinity at the onset of amphibole crystallization. Whether or not there is a steep change in crystallinity as a result of amphibole crystallization will be determined by factors such as melt composition, water

10.1029/2023GC011234

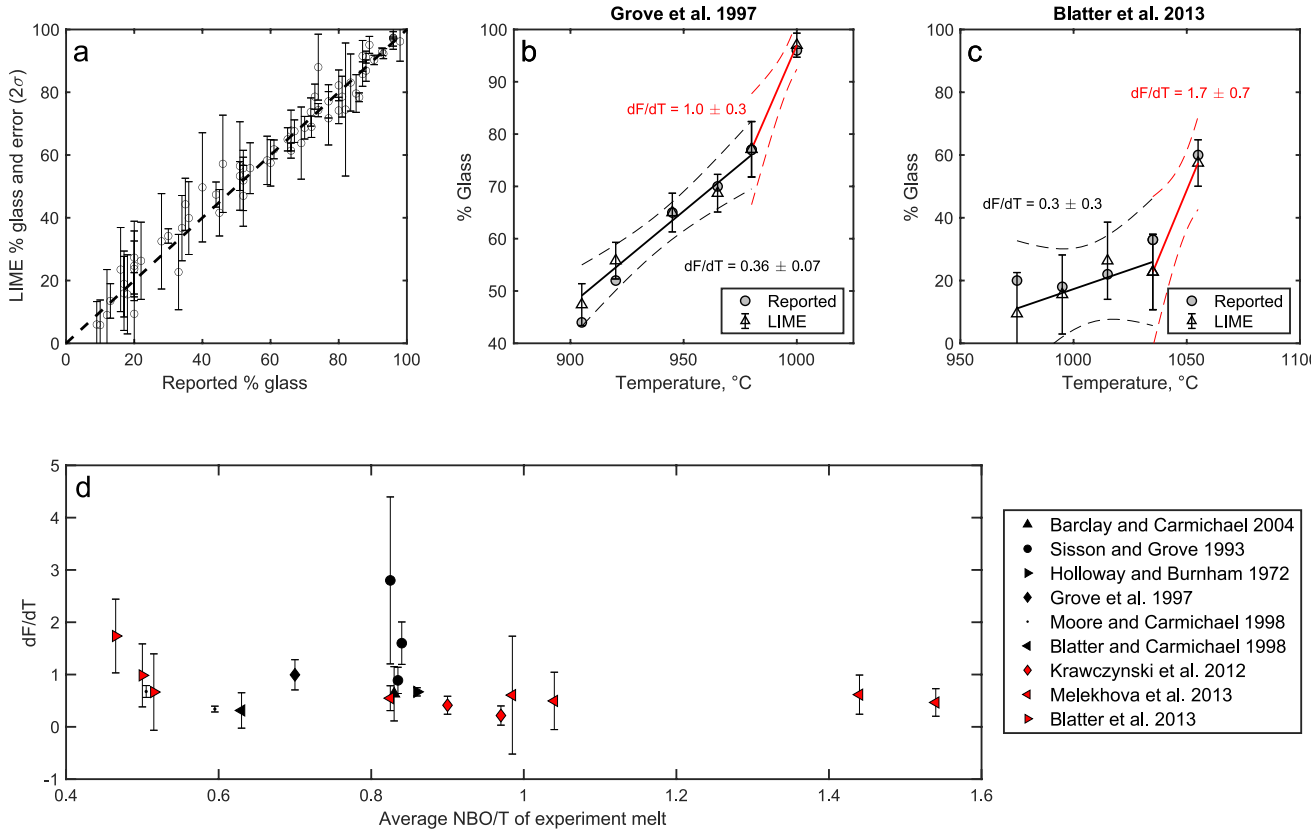

Figure 4. (a) Comparison of reported and calculated glass proportions for the set of experiments ($n = 64$) included in the magma crystallinity case study (Section 4.1). Error bars represent the 2σ error on the glass proportion calculated using LIME. The dashed line is a 1:1 line that illustrates the overall agreement between the reported and calculated glass proportions. (b and c) Glass proportion plotted as a function of temperature for two experimental series selected from Grove et al. (1997) (b) and Blatter et al. (2013) (c). For each panel, the reported values are plotted as gray circles and the LIME calculated values are depicted as triangles with 2σ error bars. The solid line corresponds to the weighted linear regression through the LIME calculated values for the amphibole-in interval (red) and amphibole crystallization (black), with the dashed lines for each fit representing the 95% confidence interval on the regression. The calculated slope and uncertainty (2σ) for each regression are given in the corresponding color. For the *dF*/*dT* reported here, *F* denotes melt percent rather than melt fraction. (d) Calculated *dF*/*dT* for the amphibole-in interval of each experimental series included in the magma crystallinity case study plotted against the average NBO/*T* for the experimental glass compositions in a given series. The data symbol reflects the experimental study from which the *dF*/*dT* was calculated (as given in the legend). Black symbols represent those studies included in the compilation of Barclay and Carmichael (2004), whereas red symbols denote experiments that have been added to the compilation by this study for comparison. The error bars depict the 2σ uncertainty on the slope as calculated from a weighted linear regression through the LIME glass proportions and experiment temperatures.

content, and pressure. A large *dF*/*dT* is expected for lower melt SiO_2, lower H_2O, and lower pressure (Barclay & Carmichael, 2004; Melekhova et al., 2013, 2015). Additionally, fractional crystallization sequences will have a smaller *dF*/*dT* because the amphibole-producing reactions between melt and earlier crystallized silicates will be suppressed (Nandedkar et al., 2014).

Using LIME, we have determined the *dF*/*dT* at the onset of amphibole crystallization for 19 amphibole-bearing experimental sequences (Figure 4). For each experiment, the proportions of glass and mineral phases were calculated from the reported bulk composition and phase compositions. The glass proportions calculated with LIME agree with the previously reported values and provide estimates for the uncertainty on the melt fraction for each experiment (Figure 4a). The *dF*/*dT* for each experimental sequence was determined from a weighted linear regression of the calculated glass proportions and the reported experimental temperatures. In many cases, the experimental sequence exhibits a non-linear trend in melt fraction with respect to temperature. For this reason, *dF*/*dT* was calculated at the "amphibole-in" interval, defined by the first amphibole-bearing experiment and the higher-temperature experiment that precedes it. When an experimental sequence contained multiple amphibole-bearing experiments, the *dF*/*dT* was also calculated for all experiments with amphibole in the crystallizing assemblage. Comparing the *dF*/*dT* for the amphibole-in interval with that of the amphibole-bearing experiments, we have determined whether the change in *dF*/*dT* associated with the onset of amphibole crystal-

lization is significant given the estimated uncertainties on the glass proportions (Figures 4b and 4c). A significant difference in slope was found for four of the 10 experimental sequences with multiple amphibole-bearing experiments. For example, the experimental series in Grove et al. (1997) and Blatter et al. (2013) have a change in slope (dF/dT) at the onset of amphibole crystallization that is significant within the estimated uncertainty (Figures 4b and 4c). For the two experiments at lower temperature (higher crystallinity) in Grove et al. (1997), the reported glass proportions are slightly different from those calculated by LIME, yet still within the estimated 2σ uncertainty (Figure 4b). In the Blatter et al. (2013) series, the reported glass proportion increases when temperature decreases from 995 to 975°C. As glass proportion is generally expected to decrease as temperature decreases, the authors attributed this supposed increase to the uncertainty of the least-squares approach to calculating modal abundances. Here we show that the average glass proportion calculated by LIME decreases in this temperature interval as expected, and the originally reported value is within the estimated 2σ uncertainty (Figure 4c).

To explore compositional effects on dF/dT, we have plotted the calculated slopes and uncertainties against the ratio of non-bridging oxygens to tetrahedral cations (NBO/T) in the experimental melts, incorporating hydrogen as a network modifier into the calculations using the reported H_2O values for the experiments (Figure 4d). Because the composition of the experimental melt differs on either side of the amphibole-in boundary, the NBO/T used for the plot in Figure 4d is the average NBO/T of the two melts. There is not a clear relationship between dF/dT and NBO/T. For water-saturated experiments conducted at 200 MPa and originally presented as a compilation in Barclay and Carmichael (2004), experimental melts with higher NBO/T generally have a higher dF/dT. In contrast, dF/dT decreases with an increase in NBO/T for the Blatter et al. (2013) series of experiments conducted on the same composition at different pressures. Though not conclusive, the data do support the theory that the onset of amphibole crystallization causes an increase in crystallinity as shown by the break in slope (dF/dT). However, the magnitude of this change is still uncertain and requires further study on the effects of water, melt composition, and pressure. Importantly, this case study illustrates the utility of LIME in addressing petrologic questions such as those pertaining to magma crystallinity.

4.2. Case Study 2: Melting Reaction Coefficients

The mass proportion method for calculating melting reactions uses the phase abundances from experimental samples to define melting reaction coefficients (e.g., Kinzler & Grove, 1992; Walter, 1998; Walter et al., 1995). The resulting reaction coefficients are used in forward melting models, as well as to model trace element evolution particularly at low degrees of melting (e.g., Jennings & Holland, 2015; Longhi, 2002; Walter et al., 1995). Experimentally determined melting reactions describe the nature of melting (i.e., peritectic or eutectic) and identify pressure-dependent changes in melting behavior.

The uncertainty on melting reaction coefficients will be determined by the uncertainty on the calculated experimental phase abundances used to define the reactions. For experiments containing olivine and orthopyroxene, calculating modal abundances by linear regression commonly yields a negative proportion for one of the silicate phases (Section 3). Using LIME to define the phase abundances for these reactions provides an estimate of uncertainty and ensures that the phase proportions are positive values.

Garnet melting with a peritectic reaction of olivine + clinopyroxene + garnet = orthopyroxene + liquid is a well-studied example of such a system (Grove et al., 2013, and references therein). We have used this reaction here as a case study for defining reaction coefficients with LIME and the electron microprobe data and starting material bulk composition reported in Walter (1998) (Figure 5). We have recalculated the phase proportions and melting reactions for a suite of experiments and compare them to the original results of Walter (1998). Walter (1998) investigated four different experimental pressures and reports a series of experiments at each pressure. We have taken the two experiments that bracket the garnet-out transition at each experimental pressure, and then calculated the reaction coefficients following the mass proportion method (Walter et al., 1995). The phase abundance uncertainties were propagated through the reaction coefficient calculation in order to determine the uncertainty on each coefficient. In order to do this calculation, the raw electron microprobe data is necessary, however, the temperature interval between the two experiments will also influence the calculated coefficients. Walter (1998) reports reaction coefficients adjusted to certain melting intervals, and thus, in order to make a direct comparison we have recalculated reaction coefficients from the phase modal abundances reported in Table 8 of Walter (1998) (Figure 6).

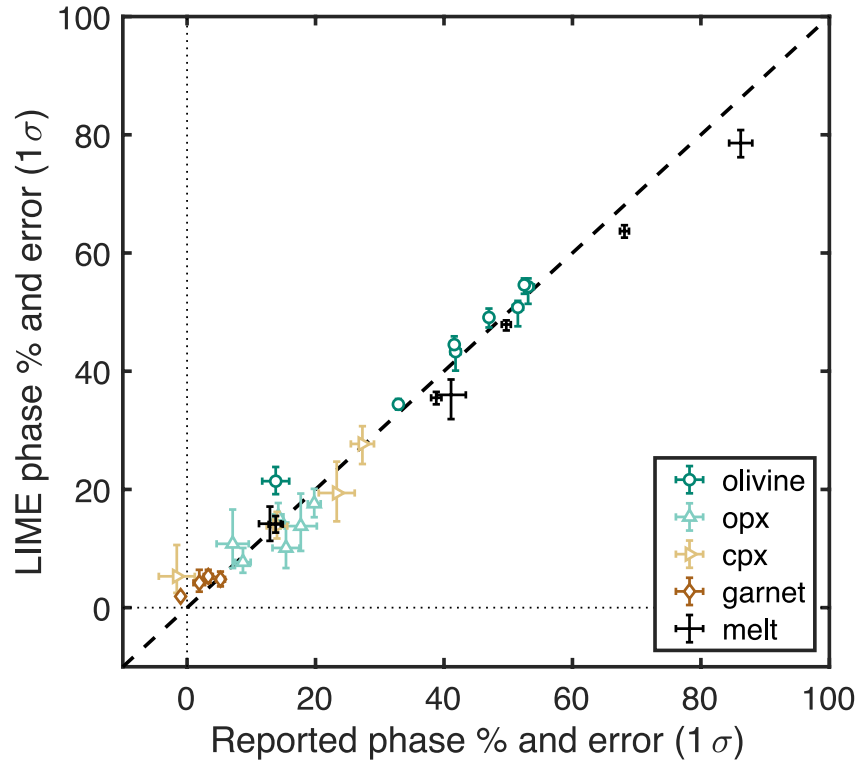

Figure 5. Comparison of reported and calculated phase proportions for the experiments bracketing the garnet-out melting reactions investigated by Walter (1998) (Section 4.2). The error bars on the values calculated from LIME represent the 15.9 and 84.1 percentiles, which, if the distribution is Gaussian, represents a 1σ error on the phase proportion (Section 3). Walter (1998) used Monte Carlo simulations to assess uncertainty in calculated phase proportions. The error bars on the reported values from Walter (1998) represent the reported 1σ standard variation. The dashed line is a 1:1 line that illustrates the overall agreement between the reported and calculated phase proportions.

The phase proportions calculated with LIME broadly agree with those reported in Walter (1998) (Figure 5). However, Walter (1998) reported negative phase proportions for two experiments (one clinopyroxene, one garnet), whereas the proportions are non-negative when calculated using LIME. This difference is non-trivial in practice, as a reaction coefficient calculated from a negative phase abundance would result in that phase appearing on the opposite side of the melting reaction. For instance, the coefficient for garnet in a garnet-out melting reaction should be negative because the phase is consumed. However, the previously reported negative phase proportion of garnet at 7 GPa results in a calculated reaction coefficient that is positive (Figure 6).

The calculated reaction coefficients reveal how the garnet melting reaction changes with pressure (Figure 6). At pressures below 7 GPa, the olivine melting reaction coefficient is near zero, indicating that olivine does not have a strong control on the reaction. However, at 7 GPa neither orthopyroxene nor clinopyroxene is present in the melting reaction, and thus olivine becomes the phase consumed in addition to garnet. At pressures below 6 GPa, clinopyroxene is consumed (negative reaction coefficient) and orthopyroxene is precipitated (positive reaction coefficient). At 6 GPa, orthopyroxene is consumed, and the coefficient calculated using the phase proportions from LIME indicates clinopyroxene is also consumed. In contrast, the clinopyroxene reaction coefficient calculated from the reported phase proportions in Walter (1998) is positive. The positive clinopyroxene reaction coefficient does not indicate a change in melting behavior but is instead an artifact of the negative reported modal abundance for clinopyroxene.

5. Conclusion

LIME is a practical log-ratio approach to mass-balance calculations in igneous and metamorphic systems that provides benefits over traditional multiple linear regression techniques. When compositions are mapped from the

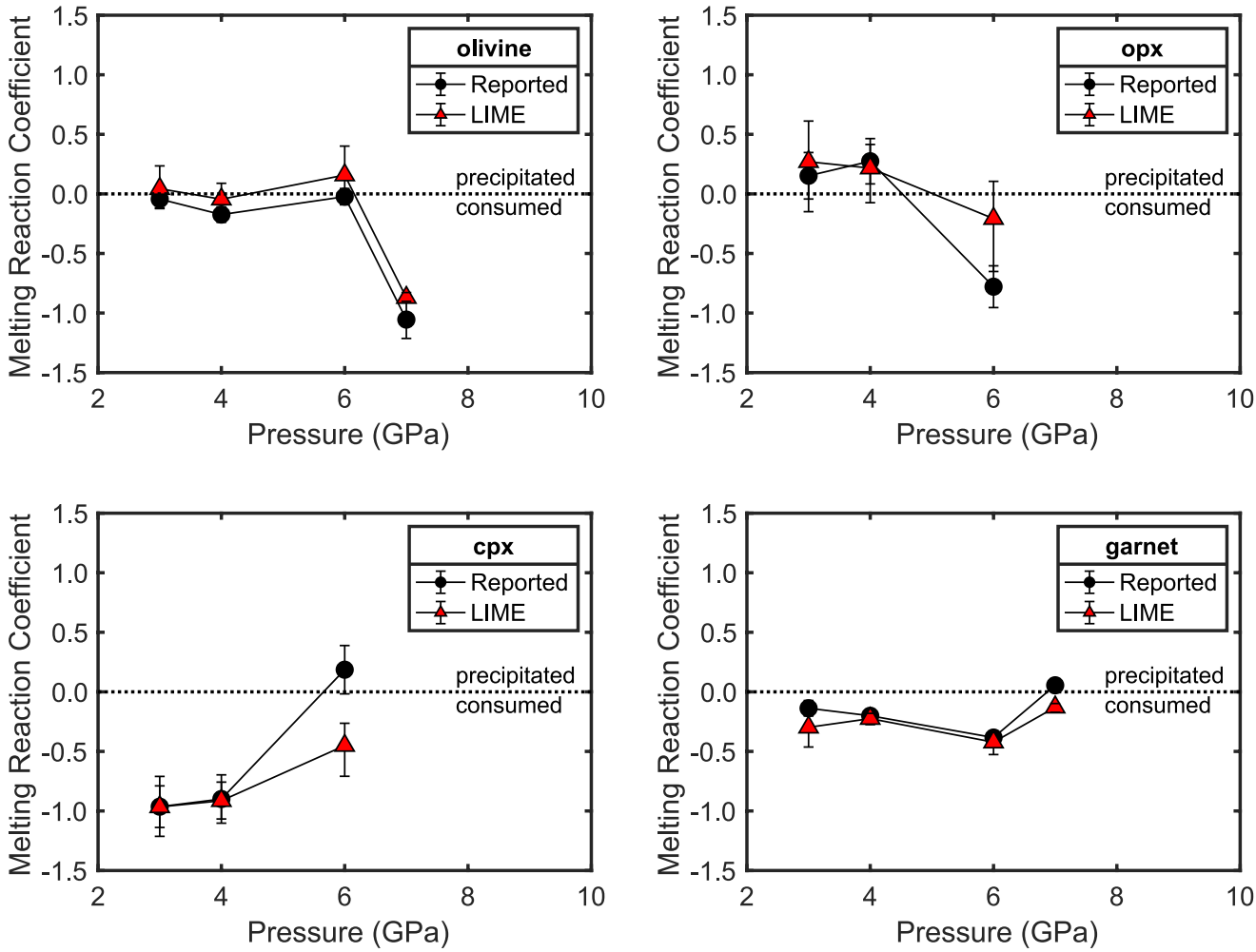

Figure 6. Garnet-out melting reaction coefficients calculated from the phase abundances in experiments from Walter (1998) (Section 4.2, Figure 5). Coefficients are plotted as a function of pressure and each phase has a separate plot ("opx" = orthopyroxene, "cpx" = clinopyroxene). The uncertainties on phase modal proportions have been propagated through the reaction coefficient calculation and the plotted error bars represent the 1σ error.

inverse log-ratio space back to composition space, the algorithm reports solutions with realistic uncertainty on end-member abundances. In using LIME to determine phase abundances, we:

- Estimate phase abundances based on user-input compositional uncertainties.
- Report accurate uncertainties, which are often asymmetric.
- Propagate uncertainty in phase proportions through subsequent calculations of magma crystallinity and melting reaction coefficients.

We have demonstrated herein the ability of LIME to provide accurate confidence estimates for fundamental quantitative calculations in petrology, and this utility can be extended to a variety of petrologic applications. The LIME algorithm can also be used to reassess cases where previously reported data yielded negative phase abundances or did not include estimates of uncertainty. Looking forward, we hope that the LIME approach may be established as a new standard benchmark for petrologic studies that require phase abundance calculations.

Appendix A: Computing the Isometric Log-Ratio Transform

The ilr-transform of a D-component composition vector returns its projection on an orthonormal basis of S^D (see Egozcue et al. (2003) and Chave (2017) for extensive definitions and detailed computation). Such a basis can be

found by performing the Gram-Schmidt procedure on a known-set of $D-1$ independent compositions in S^D. This calculation is made especially convenient by the use of Aitchison's centered log-ratio transform (clr), defined by

$$clr(x)_i = \log\left[\frac{x_i}{g(x)}\right], \tag{A1}$$

where $g(x)$ denotes the geometric mean of vector x, that is,

$$g(x) = \left[\prod_{i=1}^{D} x_i\right]^{\frac{1}{D}}. \tag{A2}$$

Note that the inverse-clr transform (clr^{-1}) is easily computed by taking the exponential of each component of $\text{clr}(x)$, and then renormalizing them so they add up to one.

Once the clr of composition x is known, we need to define a simple orthonormal (in the Euclidian sense) basis of a $(D-1)$ dimensional subspace of $\mathbb{R}^D$, termed $\mathbb{B}$. A convenient one is the set of $(D-1)$ unit vectors $u_i \in \mathbb{R}^D$, $i = 1, 2, \ldots, D-1$ defined by

$$u_i = \sqrt{\frac{i}{i+1}}\left[\underbrace{\frac{1}{i} \ \ldots \ \frac{1}{i}}_{\text{i times}} \ -1 \ 0 \ \ldots \ 0\right] \tag{A3}$$

We then store these vectors as rows of a $(D-1) \times D$ matrix U, and obtain the following relation between the ilr- and clr- transforms:

$$ilr(x) = clr(x)\,U^T \tag{A4}$$

and

$$clr(x) = ilr(x)\,U \tag{A5}$$

Other transition matrices U can be obtained by choosing a different orthonormal basis of $\mathbb{B}$. Ad-hoc choices (for detailed compositional inference studies) could be based on sequential binary partitions that make physical sense, but these are beyond the scope of this contribution.

Appendix B: Eliminating Zeroes in Petrologic Data Sets

Because the logarithm of zero is undefined, values of zero cannot serve as inputs to the log-ratio transform, which is the critical part of the LIME algorithm. Compositional data in petrology can include values of "zero" for multiple reasons. For instance, when an element is not analyzed, such as Al_2O_3 in plagioclase, or when the concentration of an element is below the analytical detection limit and an author chose to represent this as a zero. To make data that include zeros suitable for LIME, we have developed a normative-type calculation that recomputes phase compositions in a way that eliminates zeroes.

We accomplished this using a balanced approach of amalgamation and replacement of zero values. Amalgamation involves the addition of two or more oxides to make a parameter for fitting that is a combination of the two. This reduces the amount of available data to make a fit and can make the system less over-determined. However, this amalgamation often has the benefit of preventing the algorithm's iterations from getting stuck in local minima and over-emphasizing the importance of elements whose abundances are poorly constrained (i.e., near detection limit values).

There are times when amalgamation must be avoided in order to preserve data that are critical to the mass balance. These cases primarily include calculations where an element resides in a single phase at high abundance and is rare or absent in the other phases. A good example of this is when a phosphate, such as apatite, is present in an assemblage. In this case, phosphorus is a critical element in determining the abundance of the phosphate phase, so amalgamating the phosphorus data with another element will significantly increase

the uncertainty in the calculated phosphate abundance. In cases such as these, we have chosen not to remove zeroes by amalgamation, and instead replace zeroes with a trivial amount of an element (usually 0.01 wt%). In this way, the minor amount of the element in the phases with zero abundance does not affect the overall mass balance in an appreciable way, but the element can still be used in the LIME algorithm to mass balance the phase with high abundance. The cut-off values for whether zero data get amalgamated or replaced have been specifically defined with certain minerals in mind, and these minerals are noted within each amalgamation step description.

Below, we detail the normative calculation procedures implemented by LIME to help the user understand the data processing and the meaning of the amalgamated output:

To begin, LIME converts all input compositions to wt.% monocation oxides and converts all FeO and Fe_2O_3 to FeO_{total}. At this point all compositions are normalized to 100 wt.%. All bulk compositions of igneous and metamorphic rocks will have at least the following oxides: SiO_2, CaO, MgO, $AlO_{1.5}$. These oxides comprise the CMAS system, which is a basic building block in natural magmatic systems and commonly used in experimental systems (e.g., haplobasalt). *Bulk compositions in LIME should always have at least these four oxides*. Following the assumption that a bulk composition contains CMAS, the subsequent steps provide a framework for the remaining relevant rock-forming elements.

Step 1. LIME eliminates oxides with zero values in the bulk composition. If an oxide absent from the bulk has a reported abundance <5 wt.% in a phase, the oxide will be removed from the end-member phase composition and the composition will be renormalized to 100 wt.% without that particular oxide. If there is a phase with ≥5 wt.% of an oxide that has a zero value in the bulk composition, that phase will not be included in the mass-balance calculation (and will be assumed of trivial abundance). If there exists a phase (e.g., chromite) for which an important oxide is missing from the input bulk composition (e.g., $CrO_{1.5}$), this suggests that the input bulk composition is not an accurate representation of the bulk sample. In this case, we suggest modifying the bulk composition (i.e., include a non-zero estimate) so that the phase of interest is not excluded from the LIME calculation.

Step 2. For any compositions which report $HO_{0.5}$, F, Cl, or NiO as compositional components, LIME will remove the components from the composition (bulk or end-member phase) and renormalize the composition to 100 wt.%.

Step 3. LIME assesses $AlO_{1.5}$ and TiO_2 for potential amalgamation with SiO_2. For $AlO_{1.5}$, if an end-member phase is found to contain >19.4 wt.% (albite), then $AlO_{1.5}$ will not be amalgamated and 0.01 wt.% will be substituted for any zero values of $AlO_{1.5}$. For TiO_2, if there is an end-member phase with >8 wt.% (magnetite-ülvospinel solid solution), TiO_2 will not be amalgamated and zero values will be replaced by 0.01 wt.%. If the reported abundance of $AlO_{1.5}$ and/or TiO_2 in all end-member phases is less than these pre-defined values (including zero), the oxide(s) will be amalgamated with SiO_2 for the bulk and phase compositions. For all amalgamation calculations, the command window will display the result for the user (i.e., `Amalgamating Si and Al` or `substituting zero Ti by 0.01% Ti`). Oxide labels will also be adjusted to reflect any amalgamation (i.e., "SiO_2" becomes "SiO_2-Al_2O_3-TiO_2" if both Al and Ti are amalgamated).

Step 4. LIME inspects $CrO_{1.5}$ and MnO with a method similar to that described in Step 3. If there exists an end-member phase with >4 wt.% $CrO_{1.5}$ or MnO (e.g., Cr- or Mn-bearing spinels), then a value of 0.01 wt.% will be substituted in all phases with zero reported $CrO_{1.5}$ or MnO. Otherwise, if no end-member phase contains >4 wt.% $CrO_{1.5}$ or MnO, the amount of these oxides will be set to zero and the compositions will be renormalized to 100 wt.%.

Step 5. If there is no MgO reported for a phase, LIME will substitute a value of 0.01 wt.%.

Step 6. If one end-member phase has no FeO reported, LIME will add the amount of FeO to MgO in all phases, thereby amalgamating FeO and MgO.

Step 7. If there is no CaO reported for an end-member phase, LIME will substitute a value of 0.01 wt.%.

Step 8. If there is no $NaO_{0.5}$ reported for an end-member phase, LIME will amalgamate $NaO_{0.5}$ with CaO in all phases.

Step 9. If an end-member phase contains >6 wt.% $KO_{0.5}$, LIME will substitute a value of 0.01 wt.% $KO_{0.5}$ for any phase with zero $KO_{0.5}$. Otherwise, $KO_{0.5}$ will be set to zero in all phases and the compositions will be renormalized.

Step 10. If there is a phase with >42.2 wt.% $PO_{2.5}$ (fluorapatite), LIME will replace any zero reported $PO_{2.5}$ in an end-member phase with 0.01 wt.%. If there is no phase with >42.2 wt.% $PO_{2.5}$, LIME will remove $PO_{2.5}$ from all phases and renormalize the compositions.

Appendix C: Assembling Covariance Matrices for Compositional Uncertainty

All the procedures described below assume zeroes have been eliminated from the data set following the method of Appendix B. At this stage, the data set should contain P mineral phases and C possibly amalgamated components (oxides).

C1. Covariance Matrix for the Bulk Composition

Uncertainties on bulk compositions are rarely reported in the experimental literature. When reported, the bulk composition uncertainties can be directly incorporated into LIME (Appendix E). When bulk uncertainty is not available, we assume that each component of the bulk composition vector is known within the same relative uncertainty (e.g., 1%, user-specified input parameter on Line 10 of `lime_input.m`). We then use the bulk uncertainty (either reported or estimated) to draw 1,000 random bulk compositions, which we transform into their ilr counterparts. The variance of each component of the ilr-transformed vectors is then used to populate the diagonal of matrix C_D, which is of size $(C-1)\times(C-1)$.

C2. Covariance Matrix for End-Member Compositions

The composition of end-member phases is typically determined from several geochemical analyses conducted on multiple grains or at various locations within a single grain. We take advantage of this redundancy and compute the ilr-transform of each individual measurement (a C-element composition vector). The resulting list of ilr-transformed vectors is used to compute a $(C-1)\times(C-1)$ covariance matrix $C_E^{(i)}$, which includes off-diagonal components describing possible covariation among certain oxides.

If only one compositional measurement exists for an end-member phase, we recommend generating a small number (~10) of synthetic measurements assuming a known relative uncertainty (see `draw_compos.m` in Appendix E), and use these to determine $C_E^{(i)}$, as is systematically done for bulk composition uncertainty. This produces a diagonal matrix $C_E^{(i)}$, which does not account for possible covariations between oxide compositions within each end-member phase (e.g., SiO_2 and Al_2O_3 in plagioclase). Once a covariance matrix has been determined for each end-member composition, each $C_E^{(i)}$ is assembled into a block-diagonal matrix C_E of size $P(C-1)\times P(C-1)$.

C3. Complete Covariance Matrix for End-Member Compositions and Phase Proportions

Our algorithm iterates on a model vector m that concatenates the ilr of the phase proportions and that of the end-member compositions. Its associated (prior) covariance matrix C_{PRIOR} must therefore reflect this structure. We use C_E as the upper-left block of C_{PRIOR}. The lower-right block is simply a $(P-1)\times(P-1)$ diagonal matrix with constant coefficients set equal to 1 (or a large number).

Finally, if the end-member data is such that C_E is poorly conditioned (conditioning number below 10^{-15}), we reduce C_{PRIOR} to its diagonal to ensure its inverse is easily computed. When this occurs, the following text will display in the command window: `WARNING - end-member covariance matrix badly conditioned - was reduced to its diagonal.`

Appendix D: Synthetic Tests

Inverting synthetic data sets is a straightforward way to ensure that LIME yields accurate results across a wide range of compositions. Synthetic tests also assess the limits of the LIME approach and its handling of uncertainties. To create synthetic data, we calculated mixtures of the glass, olivine, and plagioclase compositions reported by Draper and Johnston (1992). We generated 900 synthetic bulk compositions (Equation 1), each corresponding to a three-end-member mixture uniformly sampled in the corresponding (2-D) ilr-transformed space.

To incorporate compositional variation for each end-member (glass, olivine, and plagioclase), we generated 15 end-member compositions drawn from a normal distribution centered on the reported composition and a defined standard deviation of either 0.1%, 1%, or 5% relative to the mean for each oxide. These synthetic compositions are meant to represent individual analyses of a given mineral phase, which will inherently have uncertainty and spatial variability. This table of end-member compositions, along with the bulk composition, were then used as inputs for the LIME algorithm.

For each of the 900 synthetic experiments, we compared the LIME-estimated mineral proportions to the true proportion used when generating the synthetic mixture (Figure D1). Uncertainty or variability in the end-member composition is the strongest control on LIME's performance. End-member compositions are precisely known (0.1% uncertainty: Figures D1a–D1c), LIME recovers the true phase proportions almost perfectly, producing results with small error bars. When the oxide composition of individual end-members varies within 5%, LIME reflects this variability by attributing large posterior uncertainties to the phase proportions (large error bars in Figures D1g–D1i). Large uncertainties may also skew the estimate in non-straightforward ways that are specific to the chemistry of each mineral phase. As an example, in this synthetic test large end-member uncertainties make LIME underestimate large proportions of glass (Figure D1g), overestimate low proportions of plagioclase (Figure D1i), and leave olivine relatively unaffected (Figure D1h). For end-member uncertainties that

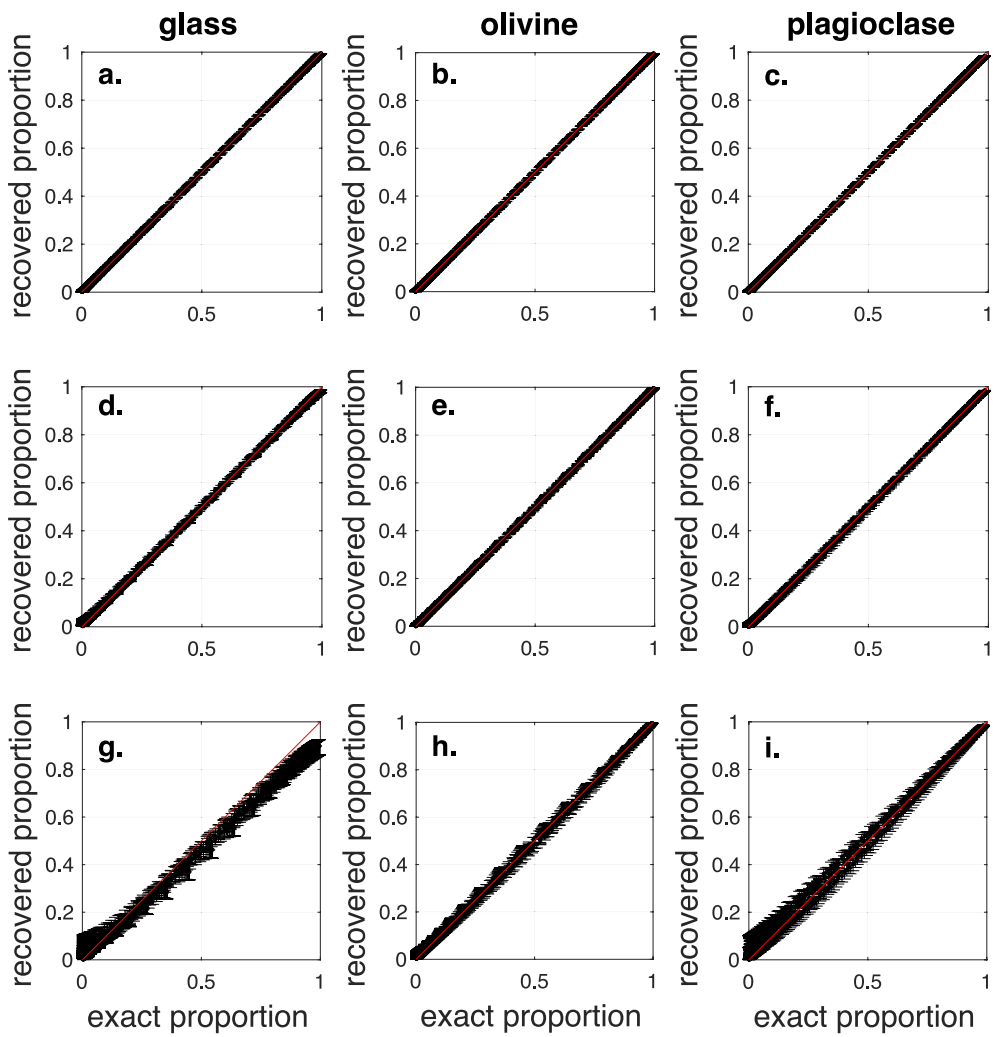

Figure D1. Recovered versus true mineral proportions from 900 synthetic experiments constructed as mixtures of glass, olivine and plagioclase. 1:1 line is plotted in red. Error bars represent 25th and 75th percentile of LIME estimates. Each row corresponds to a different assumption on the (relative) uncertainty of the compositions of glass, olivine, and plagioclase: (a–c) 0.1%; (d–f) 1%; (g–i) 5%.

are representative of typical experimental data (1%: Figures D1d–D1f), LIME recovers known phase mixtures with posterior uncertainties (error bar width) on the order of percents. From these synthetic tests, we conclude that LIME recovers balanced mixtures more accurately than mixtures where one or two end-members are in very small proportions (e.g., see the increased deviation from the red line and width of error bars close to 0 and 1 in Figures D1d–D1f).

Appendix E: User Guide

Log-ratio Inversion of Mixed End-members (LIME) is a MATLAB (The MathWorks Inc, 2022) program for calculating the phase proportions in a rock given a bulk composition and the composition of each phase. The mathematical details, outputs, and example applications are fully described in the main text. Here, instructions are presented for performing calculations within LIME.

E1. Getting Started

The plotting commands and statistical functions in the source code for LIME require that MATLAB version 2019a or later, including the "Statistics and Machine Learning Toolbox," be installed on your machine in order to perform a LIME calculation. To begin using LIME, download `LIME.zip` from the GitHub repository (github.com/kprissel/LIME). Add the provided files to your MATLAB path.

Contents of `LIME.zip`:

`lime_input.m`: This is the front-end, user-interfacing file for computing with LIME. User inputs and preferences are given, and a calculation can be completed from start to finish by running this script.

`Inputs (folder)`: Example input files for compositional data including.xlsx and.txt formats. Use these files as templates for entering your own data into LIME.

`main_script.m`: Serves as the main engine for the code by calling upon subscripts. Called in `lime_input.m` for each input composition.

`preplime.m`: Function that prepares user-input data for the formatting expected in subsequent scripts. This allows the user to input the compositional data with any oxide order and include phase labels that will be used in LIME figures.

`clr.m`: Function that computes the centered log-ratio transform from the D-dimensional S^D simplex to $\mathbb{R}^D$ (Aitchison, 1982).

`clrinv.m`: Function that computes the inverse centered log-ratio transform from $\mathbb{R}^D$ to the D-dimensional S^D simplex (Aitchison, 1982).

`draw_compos.m`: Function that generates a set of random compositions with a specified mean and standard deviation.

`forwardilr.m`: Function that computes the forward composition problem in ilr formulation.

`ilr.m`: Function that computes the isometric log-ratio transform from the D-dimensional SD simplex to $\mathbb{R}^{D-1}$ (Egozcue et al., 2003).

`ilrinv.m`: Function that computers the inverse-isometric log-ratio transform from $\mathbb{R}^{D-1}$ to the D-dimensional S^D simplex (Egozcue et al., 2003).

`lime.m`: Function that computes the log-ratio-inversion of mixed end-members.

`OKnorm.m`: Function that amalgamates oxides in a pre-defined way, as described in the manuscript text (Section 2.2).

E2. Data Inputs Required for LIME

Data required for LIME are the average bulk composition and the compositions of each phase, all reported as weight percent oxides (SiO_2, TiO_2, Al_2O_3, Fe_2O_3, Cr_2O_3, FeO, MnO, MgO, NiO, CaO, Na_2O, K_2O, P_2O_5, H_2O).

Uncertainty on the bulk composition can be input as generalized percent or using specified values for each oxide. The uncertainty on each phase composition will be determined based on the standard deviation of all the compositions input for a given phase. Because of this, you must give at least two compositional analyses for each phase. For data that has been published as averages and standard deviations of *n* points, you can reproduce a synthetic data set with those statistics using `draw_compos.m`. To do this, run the function `draw_compos(N, mean, std)`, where `N` is the number of analyses to be generated, `mean` is the average of those analyses, and `std` is the standard deviation of the normal distribution that will be used to randomly generate compositions. Note that for the compositional inputs, FeO and Fe_2O_3 get amalgamated to total FeO (FeO_T), so the phase abundance calculation will not directly include variations in Fe oxidation state. Though NiO and H_2O are accepted as inputs, the current version of LIME zeroes these oxides out prior to performing a calculation. Data should be appropriately quality checked and filtered before being input to LIME. If there is a compositional input with a total less than 95 wt.%, a warning will be displayed in the command window.

Data are input as either a spreadsheet or text file. There are multiple input file examples included in the `Inputs` folder. We suggest using `TEMPLATE.xlsx` to get started with your own data. The format of the input files has columns for the phase label and oxides. In this version of LIME, the phase label must be the first column but the oxides can be given in any order. After the row of column labels, each subsequent row corresponds to a composition.

The bulk composition is listed first in `TEMPLATE.xlsx`, and in this version of LIME the bulk composition must be given the "bulk" phase label. If the uncertainty (wt.%) is known for each oxide in "bulk," include these values in a subsequent row labeled "*U*." If no "*U*" labeled row is given for uncertainty on the bulk composition, the `pct_err_bulk` given in `lime_input.m` will be used instead (1% uncertainty default).

The remaining rows are compositional analyses for the phases, with the phase label corresponding to each row given in the first column (e.g., "phase1"). During a routine, the subfile `preplime.m` will find the "bulk" and "*U*" entries, rearrange the oxide columns to be in the expected order for the rest of the calculations, and group the rows of analyses according to the phase labels. Be careful to use the exact same label (matching characters and title case) for each row that is meant to represent the same phase. For example, if phases are labeled "clinopyroxene" and one of the clinopyroxene rows is given as "cpx," the code will treat these are separate phases. Another common issue when users first start with LIME is the oxide column labels. Be sure to use the oxide labels given in the example input files (e.g., "SiO_2"); any additional spaces or characters in the oxide column names will prevent the code from finding the expected oxide (e.g., "wt% SiO_2" or "SiO_2").

E3. Running the Code

Aside from the input files, LIME users will only need to interface with `lime_input.m` to successfully perform a calculation. Once the input file is prepared, open `lime_input.m` in MATLAB. Give the file path for the input file in the `filename` variable at Line 7. Multiple file names can be given, with each file name string separated by a comma. When running multiple input files at once, we recommend saving the outputs (set `ouput_yn` = 1 on Line 11), because variable names will be overwritten with each code iteration.

User-Specified Parameters, *Lines 9–12*: Here you can define a percent uncertainty in the bulk composition (Line 10), choose whether to save the result outputs as a separate file (Line 11), and choose whether to display the result convergence plot to assess the performance of the code (Line 12).

Additional Parameters, *Lines 19–28*: These parameters relate to the ilr conversion and calculations in ilr-space. They also define how many synthetic compositions to generate for the bulk composition (Line 23) and how many samples to generate for the posterior distribution (Line 28). The specific details of these calculations are given in Section 2.3.

The variable `NEWT.error_measure` defines how the reported error will be calculated. When this variable is set to "1," the reported errors on the posterior distribution will be the 1-sigma equivalent of a Gaussian distribution (15.9 and 84.1 percentiles). If the variable is set to "2," the errors will be the 2-sigma error equivalent for a Gaussian distribution (2.3 and 97.7 percentiles). Because not all posterior distributions have a Gaussian distribution (see asymmetry in Section 3), you can also define a vector of two percentile values to be used in the error calculation instead. If a vector of two values is specified, those percentiles will be used instead of the default 15.9

and 84.1. The provided `lime_input.m` file sets the reported errors to be the 25th and 75th percentiles of the posterior distribution.

E4. Results and Outputs

Results are reported in the command window (Figure E1) and as probability density functions in MATLAB Figure 1. If `plot_cvg` = 1 on Line 12 of `lime_input.m`, then MATLAB Figure 2 will show the convergence plot of phase proportions at each iteration of the calculation.

The command window will first display text that details the modifications made by `OKnorm.m` to the compositions during the calculation (e.g., `zero-ing oxides missing from the bulk composition` or oxide amalgamation). Then, the results are given for each phase. The input phase strings are listed as column headers in the order corresponding to the order of the reported result values. The first set of phase proportions given in the command window are the mean and standard deviation (wt.%) for each phase. The second set of phase proportions corresponds to the best fitting phase abundance (wt.%, peak center) and the wt.% units to each user-specified percentile of the PDFs shown in Figure 1. *We recommend reporting the second set of values, particularly if there is asymmetry in your results, and rounding the results to significant figures appropriate for your calculated uncertainties* (see Section 3.2). The last section of the command window result gives the recalculated and input bulk compositions, reported in the oxide components as modified by `OKnorm.m`.

If `output_yn` = 1 in Line 11 of `lime_input.m`, a folder will be created when you run the code. It will be titled the name of your input file plus "_LIME" and contain an Excel file (_LIME_results.xlsx) with a tab for the command window output ("Main Output") as well as the 80,000 modeled abundances used to generate the results ("Modeled abundances"). This folder will also contain a PDF of Figure 1 (_LIME_figure.pdf) and a duplicate file of the input compositions.

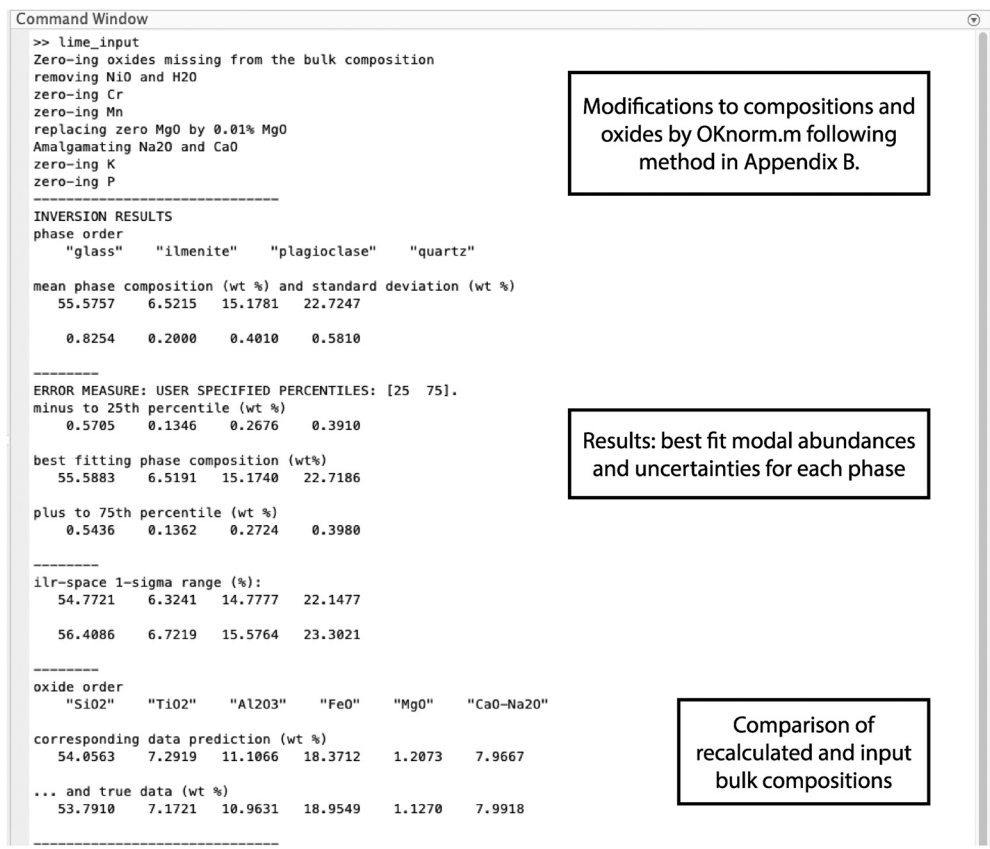

Figure E1. Screenshot of command window after successful run of `lime_input.m` with provided input file J047.xlsx.

E5. Calculating Experimental Iron Loss

Reaction between Fe-rich samples and metal sample containers is a well-known occurrence in experimental petrology (Borisov & Jones, 1999; Grove, 1982). During an experiment, Fe can diffuse out of the experimental sample and into the metal container. The amount of iron "lost" to the capsule will depend on the metal alloy composition of the container as well as the oxygen fugacity, with increased partitioning of Fe into metal at lower oxygen fugacities.

We have developed a way to estimate Fe loss using LIME. To do this, we perform a phase abundance calculation excluding Fe from the bulk and phase compositions. Using the phase proportions from the Fe-free calculation and the average phase compositions (including Fe), we then recalculate the bulk Fe composition. Comparing the recalculated bulk Fe content to the initial bulk Fe content determines the percent of Fe that was lost from the bulk sample during the experiment. When performing this calculation, the Fe-bearing and Fe-free phase abundances should be compared to assess how influential Fe content is on the calculated phase abundances. For example, an Fe-free calculation will not be reliable if an experiment contains a phase that is primarily iron (e.g., magnetite or Fe metal), because that phase will no longer be included in the phase abundance calculation. We have used this method previously to estimate Fe loss for olivine- and glass-bearing experiments (Prissel et al., 2018) as well as ilmenite- and glass-bearing experiments (Prissel, 2020). In the case of ilmenite, zeroing out Fe makes the ilmenite proportion predominantly dependent on the Ti compositions of the ilmenite, glass, and bulk composition.

In Prissel et al. (2018), we analyzed the Fe content of Re metal sample containers using both electron microprobe and inductively coupled plasma mass spectrometry. The LIME-estimated amount of Fe lost from the experimental sample was greater than the amount of Fe in found in the Re sample container. This discrepancy indicated another source of Fe loss, evaporation, which was also supported by Fe isotopic analysis of the samples and metal containers.

To calculate Fe loss with LIME, begin by inputting your data as you would for a normal calculation. To zero out Fe, go to Line 5 of the `main_script.m` file and add "0" as a third input in the "preplime" function: 0 = Fe-free, 1 = Fe-bearing (This optional input is used for the loop at Line 48–54 in `preplime.m`.) With "0" as the third input in the preplime function (i.e., `preplime(filename{f},pct_err_bulk,0)`), running `lime_input.m` will conduct an Fe-free phase abundance calculation. Then, the resulting proportions from the Fe-free calculation can be used to determine bulk Fe loss as described above.

Data Availability Statement

The initial version of the source code used for this manuscript, as well as the example input files, have been provided with DOI on Zenodo (Prissel et al., 2023) and future versions of the code will be maintained on GitHub (https://github.com/kprissel/LIME).

Acknowledgments
The authors extend their appreciation to the LIME users who tested the algorithm during its developmental stages, helping us to identify areas for improvement in the code and user interface. Specifically, we thank Meghan Guild, Andrea Goltz, and Gordon Moore for in-depth conversations over the course of this project. We also thank Jordan Lubbers for an insightful review of the manuscript and Paul Asimow for thoughtful editorial handling. KP acknowledges financial support from the McDonnell Center for Space Sciences at Washington University in St. Louis, the NASA Earth and Space Sciences Fellowship program, Carnegie Earth and Planets Laboratory, as well as the Jacobs JETS contract at NASA Johnson Space Center. J.-A.O. acknowledges A. Chave and the late A. Tarantola for inspiring the initial version of LIME through their 2010 and 2008 courses at WHOI and IPGP. MJK acknowledges support from NSF awards EAR-2047960, EAR-2042386, and NASA award 80NSSC20K0640.

References

Aitchison, J. (1982). The statistical analysis of compositional data. *Journal of the Royal Statistical Society: Series B*, *44*(2), 139–160. https://doi.org/10.1111/j.2517-6161.1982.tb01195.x

Aitchison, J. (1984). The statistical analysis of geochemical compositions. *Journal of the International Association for Mathematical Geology*, *16*(6), 531–564. https://doi.org/10.1007/bf01029316

Albarede, F., & Provost, A. (1977). Petrological and geochemical mass-balance equations: An algorithm for least-square fitting and general error analysis. *Computers & Geosciences*, *3*(2), 309–326. https://doi.org/10.1016/0098-3004(77)90007-3

Baker, M. B., Grove, T. L., & Price, R. (1994). Primitive basalts and andesites from the Mt. Shasta region, N. California: Products of varying melt fraction and water content. *Contributions to Mineralogy and Petrology*, *118*(2), 111–129. https://doi.org/10.1007/bf01052863

Barclay, J., & Carmichael, I. (2004). A hornblende basalt from western Mexico: Water-saturated phase relations constrain a pressure–temperature window of eruptibility. *Journal of Petrology*, *45*(3), 485–506. https://doi.org/10.1093/petrology/egg091

Blatter, D. L., Sisson, T. W., & Hankins, W. B. (2013). Crystallization of oxidized, moderately hydrous arc basalt at mid-to lower-crustal pressures: Implications for andesite genesis. *Contributions to Mineralogy and Petrology*, *166*(3), 861–886. https://doi.org/10.1007/s00410-013-0920-3

Blundy, J., & Cashman, K. (2001). Ascent-driven crystallisation of dacite magmas at Mount St. Helens, 1980–1986. *Contributions to Mineralogy and Petrology*, *140*(6), 631–650. https://doi.org/10.1007/s004100000219

Blundy, J., Cashman, K. V., & Berlo, K. (2008). Evolving magma storage conditions beneath Mount St. Helens inferred from chemical variations in melt inclusions from the 1980-1986 and current (2004-2006) eruptions (Technical Report). *US Geological Survey*, 755–790. https://doi.org/10.3133/pp175033

Borisov, A., & Jones, J. H. (1999). An evaluation of Re, as an alternative to Pt, for the 1 bar loop technique: An experimental study at 1400°C. *American Mineralogist*, *84*(10), 1528–1534. https://doi.org/10.2138/am-1999-1006

Brown, S., & Grove, T. (2015). Origin of the Apollo 14, 15, and 17 yellow ultramafic glasses by mixing of deep cumulate remelts. *Geochimica et Cosmochimica Acta*, *171*, 201–215. https://doi.org/10.1016/j.gca.2015.09.001

Bryan, W. B., Finger, L. W., & Chayes, F. (1969). Estimating proportions in petrographic mixing equations by least-squares approximation. *Science*, *163*(3870), 926–927. https://doi.org/10.1126/science.163.3870.926

Canil, D., & Bellis, A. J. (2008). Phase equilibria in a volatile-free kimberlite at 0.1 MPa and the search for primary kimberlite magma. *Lithos*, *105*(1–2), 111–117. https://doi.org/10.1016/j.lithos.2008.02.011

Cashman, K. V. (1992). Groundmass crystallization of Mount St. Helens dacite, 1980–1986: A tool for interpreting shallow magmatic processes. *Contributions to Mineralogy and Petrology*, *109*(4), 431–449. https://doi.org/10.1007/bf00306547

Cashman, K. V., & Marsh, B. D. (1988). Crystal size distribution (CSD) in rocks and the kinetics and dynamics of crystallization II: Makaopuhi lava lake. *Contributions to Mineralogy and Petrology*, *99*(3), 292–305. https://doi.org/10.1007/bf00375363

Cashman, K. V., Sparks, R. S. J., & Blundy, J. D. (2017). Vertically extensive and unstable magmatic systems: A unified view of igneous processes. *Science*, *355*(6331). https://doi.org/10.1126/science.aag3055

Chave, A. (2017). *Computational statistics in the Earth sciences with applications in MATLAB*. Cambridge University Press.

Chayes, F. (1960). On correlation between variables of constant sum. *Journal of Geophysical Research*, *65*(12), 4185–4193. https://doi.org/10.1029/jz065i012p04185

Codillo, E., Klein, F., & Marschall, H. (2022). Preferential formation of chlorite over talc during Si-metasomatism of ultramafic rocks in subduction zones. *Geophysical Research Letters*, *49*(19), e2022GL100218. https://doi.org/10.1029/2022gl100218

Collinet, M., Médard, E., Charlier, B., Vander Auwera, J., & Grove, T. L. (2015). Melting of the primitive Martian mantle at 0.5–2.2 GPa and the origin of basalts and alkaline rocks on Mars. *Earth and Planetary Science Letters*, *427*, 83–94. https://doi.org/10.1016/j.epsl.2015.06.056

Davis, F. A., & Hirschmann, M. M. (2013). The effects of K_2O on the compositions of near-solidus melts of garnet peridotite at 3 GPa and the origin of basalts from enriched mantle. *Contributions to Mineralogy and Petrology*, *166*(4), 1029–1046. https://doi.org/10.1007/s00410-013-0907-0

Dawson, H., & Krawczynski, M. J. (2022). Experimental study of halogen partitioning between amphibole and melt in subduction magmas. In *AGU fall meeting abstracts* V42G–0132.

DePaolo, D. J. (1981). Trace element and isotopic effects of combined wallrock assimilation and fractional crystallization. *Earth and Planetary Science Letters*, *53*(2), 189–202. https://doi.org/10.1016/0012-821x(81)90153-9

Draper, D. S., & Johnston, A. D. (1992). Anhydrous PT phase relations of an Aleutian high-MgO basalt: An investigation of the role of olivine-liquid reaction in the generation of arc high-alumina basalts. *Contributions to Mineralogy and Petrology*, *112*(4), 501–519. https://doi.org/10.1007/bf00310781

Egozcue, J., Pawlowsky-Glahn, V., Mateu-Figueras, G., & Barceló-Vidal, C. (2003). Isometric logratio transformations for compositional data analysis. *Mathematical Geology*, *35*(8), 279–300.

Foden, J., & Green, D. (1992). Possible role of amphibole in the origin of andesite: Some experimental and natural evidence. *Contributions to Mineralogy and Petrology*, *109*(4), 479–493. https://doi.org/10.1007/bf00306551

Gavrilenko, M., Krawczynski, M., Ruprecht, P., Li, W., & Catalano, J. G. (2019). The quench control of water estimates in convergent margin magmas. *American Mineralogist: Journal of Earth and Planetary Materials*, *104*(7), 936–948. https://doi.org/10.2138/am-2019-6735

Ghiorso, M. S. (1983). LSEQIEQ: A FORTRAN IV subroutine package for the analysis of multiple linear regression problems with possibly deficient pseudorank and linear equality and inequality constraints. *Computers & Geosciences*, *9*(3), 391–416. https://doi.org/10.1016/0098-3004(83)90008-0

Goltz, A. E., Krawczynski, M. J., McCanta, M. C., & Darby Dyar, M. (2022). Experimental calibration of an Fe^{3+}/Fe^{2+}-in-amphibole oxybarometer and its application to shallow magmatic processes at Shiveluch Volcano, Kamchatka. *American Mineralogist*, *107*(11), 2084–2100. https://doi.org/10.2138/am-2022-8031

Greenacre, M., Grunsky, E., Bacon-Shone, J., Erb, I., & Quinn, T. (2023). Aitchison's compositional data analysis 40 years on: A reappraisal. *Statistical Science*, *1*(1), 1–25. https://doi.org/10.1214/22-sts880

Grove, T., Donnelly-Nolan, J. M., & Housh, T. (1997). Magmatic processes that generated the rhyolite of Glass Mountain, Medicine Lake volcano, N. California. *Contributions to Mineralogy and Petrology*, *127*(3), 205–223. https://doi.org/10.1007/s004100050276

Grove, T. L. (1982). Use of FePt alloys to eliminate the iron loss problem in 1 atmosphere gas mixing experiments: Theoretical and practical considerations. *Contributions to Mineralogy and Petrology*, *78*(3), 298–304. https://doi.org/10.1007/BF00398924

Grove, T. L., Holbig, E. S., Barr, J. A., Till, C. B., & Krawczynski, M. J. (2013). Melts of garnet lherzolite: Experiments, models and comparison to melts of pyroxenite and carbonated lherzolite. *Contributions to Mineralogy and Petrology*, *166*(3), 887–910. https://doi.org/10.1007/s00410-013-0899-9

Guenther, M. E., Krein, S. M. B., & Grove, T. L. (2022). The influence of variable oxygen fugacity on the source depths of lunar high-titanium ultramafic glasses. *Geochimica et Cosmochimica Acta*, *334*, 217–230. https://doi.org/10.1016/j.gca.2022.07.023

Helz, R. T. (1973). Phase relations of basalts in their melting range at P_{H2O}=5 kb as a function of oxygen fugacity: Part I. Mafic phases. *Journal of Petrology*, *14*(2), 249–302. https://doi.org/10.1093/petrology/14.2.249

Holloway, J. R., & Burnham, C. W. (1972). Melting relations of basalt with equilibrium water pressure less than total pressure. *Journal of Petrology*, *13*(1), 1–29. https://doi.org/10.1093/petrology/13.1.1

Huber, C., Bachmann, O., & Dufek, J. (2012). Crystal-poor versus crystal-rich ignimbrites: A competition between stirring and reactivation. *Geology*, *40*(2), 115–118. https://doi.org/10.1130/g32425.1

Jennings, E. S., & Holland, T. J. (2015). A simple thermodynamic model for melting of peridotite in the system NCFMASOCr. *Journal of Petrology*, *56*(5), 869–892. https://doi.org/10.1093/petrology/egv020

Kinzler, R. J., & Grove, T. L. (1992). Primary magmas of mid-ocean ridge basalts 1. Experiments and methods. *Journal of Geophysical Research*, *97*(B5), 6885–6906. https://doi.org/10.1029/91jb02840

Korotev, R. L., & Kremser, D. T. (1992). Compositional variations in Apollo 17 soils and their relationship to the geology of the Taurus-Littrow site. *Proceedings of Lunar and Planetary Science*, *22*, 275–301.

Krawczynski, M. J., & Grove, T. L. (2012). Experimental investigation of the influence of oxygen fugacity on the source depths for high titanium lunar ultramafic magmas. *Geochimica et Cosmochimica Acta*, *79*, 1–19. https://doi.org/10.1016/j.gca.2011.10.043

Krawczynski, M. J., & Olive, J. L. (2011). A new fitting algorithm for petrological mass-balance problems. *AGU fall meeting abstracts* (Vol. 2011, pp. V53B–V2613).

Langmuir, C. H., Klein, E. M., & Plank, T. (1992). Petrological systematics of mid-ocean ridge basalts: Constraints on melt generation beneath ocean ridges. *Mantle Flow and Melt Generation at Mid-Ocean Ridges*, 183–280. https://doi.org/10.1029/gm071p0183

Li, X., Zhang, C., Almeev, R. R., & Holtz, F. (2020). Geobalance: An excel VBA program for mass balance calculation in geosciences. *Geochemistry*, *80*(2), 125629. https://doi.org/10.1016/j.chemer.2020.125629

Longhi, J. (2002). Some phase equilibrium systematics of lherzolite melting: I. *Geochemistry, Geophysics, Geosystems*, *3*(3), 1–33. https://doi.org/10.1029/2001gc000204

Marsh, B. (1981). On the crystallinity, probability of occurrence, and rheology of lava and magma. *Contributions to Mineralogy and Petrology*, *78*(1), 85–98. https://doi.org/10.1007/bf00371146

Melekhova, E., Annen, C., & Blundy, J. (2013). Compositional gaps in igneous rock suites controlled by magma system heat and water content. *Nature Geoscience*, *6*(5), 385–390. https://doi.org/10.1038/ngeo1781

Melekhova, E., Blundy, J., Robertson, R., & Humphreys, M. C. (2015). Experimental evidence for polybaric differentiation of primitive arc basalt beneath St. Vincent, Lesser Antilles. *Journal of Petrology*, *56*(1), 161–192. https://doi.org/10.1093/petrology/egu074

Miesch, A. T. (1969). The constant sum problem in geochemistry. In *Computer applications in the Earth sciences* (pp. 161–176). https://doi.org/10.1007/978-1-4615-8633-3_10

Mitchell, A. L., & Grove, T. L. (2015). Melting the hydrous, subarc mantle: The origin of primitive andesites. *Contributions to Mineralogy and Petrology*, *170*(2), 1–23. https://doi.org/10.1007/s00410-015-1161-4

Moore, G., & Carmichael, I. (1998). The hydrous phase equilibria (to 3 kbar) of an andesite and basaltic andesite from western Mexico: Constraints on water content and conditions of phenocryst growth. *Contributions to Mineralogy and Petrology*, *130*(3–4), 304–319. https://doi.org/10.1007/s004100050367

Nandedkar, R. H., Ulmer, P., & Müntener, O. (2014). Fractional crystallization of primitive, hydrous arc magmas: An experimental study at 0.7 GPa. *Contributions to Mineralogy and Petrology*, *167*(6), 1015. https://doi.org/10.1007/s00410-014-1015-5

Parmigiani, A., Huber, C., & Bachmann, O. (2014). Mush microphysics and the reactivation of crystal-rich magma reservoirs. *Journal of Geophysical Research: Solid Earth*, *119*(8), 6308–6322. https://doi.org/10.1002/2014jb011124

Pawlowsky-Glahn, V., Egozcue, J. J., & Tolosana-Delgado, R. (2015). *Modeling and analysis of compositional data*. John Wiley & Sons.

Pichavant, M., & Macdonald, R. (2007). Crystallization of primitive basaltic magmas at crustal pressures and genesis of the calc-alkaline igneous suite: Experimental evidence from St. Vincent, Lesser Antilles arc. *Contributions to Mineralogy and Petrology*, *154*(5), 535–558. https://doi.org/10.1007/s00410-007-0208-6

Prissel, K. (2020). Experimental constraints on igneous iron isotopic fractionation and diffusion. St. Louis Doctoral Dissertation. Washington University.

Prissel, K., Krawczynski, M. J., Nie, N. X., Dauphas, N., Couvy, H., Hu, M. Y., et al. (2018). Experimentally determined effects of olivine crystallization and melt titanium content on iron isotopic fractionation in planetary basalts. *Geochimica et Cosmochimica Acta*, *238*, 580–598. https://doi.org/10.1016/j.gca.2018.07.028

Prissel, K., Olive, J.-A., & Krawczynski, M. J. (2023). LIME: A log-ratio-based algorithm for petrologic mass-balance problems and uncertainty assessment [Software]. Zenodo. https://doi.org/10.5281/zenodo.8135998

Ray, W. H., & Szekely, J. (1973). *Process optimization, with applications in metallurgy and chemical engineering*. Wiley.

Reid, M., Gancarz, A., & Albee, A. (1973). Constrained least-squares analysis of petrologic problems with an application to lunar sample 12040. *Earth and Planetary Science Letters*, *17*(2), 433–445. https://doi.org/10.1016/0012-821x(73)90212-4

Sisson, T., & Grove, T. (1993). Experimental investigations of the role of H_2O in calc-alkaline differentiation and subduction zone magmatism. *Contributions to Mineralogy and Petrology*, *113*(2), 143–166. https://doi.org/10.1007/bf00283225

Stadermann, A. C., Jolliff, B. L., Krawczynski, M. J., Hamilton, C. W., & Barnes, J. J. (2022). Analysis and experimental investigation of Apollo sample 12032, 366-18, a chemically evolved basalt from the Moon. *Meteoritics & Planetary Sciences*, *57*(4), 794–816. https://doi.org/10.1111/maps.13795

Stormer, J. C., Jr., & Nicholls, J. (1978). XLFRAC: A program for the interactive testing of magmatic differentiation models. *Computers & Geosciences*, *4*(2), 143–159. https://doi.org/10.1016/0098-3004(78)90083-3

Tarantola, A. (1981). *Inverse problem theory*. Elsevier.

The MathWorks Inc. (2022). *Matlab version: 9.13.0 (r2022b)*. The MathWorks Inc. Retrieved from https://www.mathworks.com

Walter, M. J. (1998). Melting of garnet peridotite and the origin of komatiite and depleted lithosphere. *Journal of Petrology*, *39*(1), 29–60. https://doi.org/10.1093/petroj/39.1.29

Walter, M. J., Sisson, T. W., & Presnall, D. C. (1995). A mass proportion method for calculating melting reactions and application to melting of model upper mantle lherzolite. *Earth and Planetary Science Letters*, *135*(1–4), 77–90. https://doi.org/10.1016/0012-821x(95)00148-6

Wright, T. L., & Doherty, P. C. (1970). A linear programming and least squares computer method for solving petrologic mixing problems. *Geological Society of America Bulletin*, *81*(7), 1995–2008. https://doi.org/10.1130/0016-7606(1970)81[1995:alpals]2.0.co;2

Zhang, Y., Namur, O., & Charlier, B. (2023). Experimental study of high-Ti and low-Ti basalts: Liquid lines of descent and silicate liquid immiscibility in large igneous provinces. *Contributions to Mineralogy and Petrology*, *178*(1), 7. https://doi.org/10.1007/s00410-022-01990-x