

Appendix 1: Methods

Error calculation

In order to estimate the analytical uncertainty associated with manual tracing of crystals and subsequent image processing, we performed several experiments. Firstly, duplicate tracings by different observers were analyzed using the same images (slab + thin section) from 1 sample to generate raw grain size data (Tables A1, A2). Second, multiple acquisitions of data (5x) from the same tracing of a sample were conducted in ImageJTM to estimate the error caused by changing the threshold parameter of the analyzed image (Tables A1, A2). Thresholding is the means by which the image analysis program separates or segments particles from the background. Individual pixels in a grayscale image are marked as “object” pixels if their value is greater than some threshold value (provided an object is brighter than the background) and as “background” pixels otherwise. By incrementally changing the threshold value by steps of 5 up to +/- 10 around a default value in ImageJTM (210 on a 0 to 255 scale), the olivine crystal pixilation could be slightly enhanced or dilated, thereby changing the size and/or number of olivine crystals measured by ImageJTM. Thirdly, normalization factors (eq.A1) were varied (A_i^N) to assess variance due to necessity of rounding this value to the closest integer. Total analytical error in particle areas and frequencies are the summation of uncertainties due to tracing, thresholding, and normalization ($\sigma_a^2 = \sigma_{th}^2 + \sigma_{tr}^2 + \sigma_N^2$), and is found to be less than the variance due to sampling, such that $\sigma_a^2 < \sigma_s^2$ (e.g., Fig.5a; 5b).

Stereological corrections

Data generated from a two dimensional surface do not necessarily reflect the true sizes of the crystals in three dimensions. To account for these complications, we use a three-step data organization and correction process involving: a) defining characteristic shapes from raw 2D intersection data, using CSDSliceTM (Morgan and Jerram, 2006) to obtain best-approximations of 3D ellipse axial ratios; b) organizing crystals in bins (e.g., $\log_{0.5}$, \log_2) by area and number, and; c) inserting binned data into CSDCorrectionsTM 1.3 (Higgins, 2000) to mitigate the cut-section and intersection-probability effect (sensu Higgins, 2006) and then calculate 3D olivine crystal populations by applying correction factors for known properties (i.e. crystal shape, modal %, rock fabric,). The result is a dataset that comprises scale-normalized, scale-truncated, binned, and stereologically-corrected olivine crystal populations. Variations of the correction factors for crystal shape, modal %, rock fabric, and bin sizes create datasets shown by the shaded region in Figure 8a, and are within analytical error of sample CK05.

Shape assessment. To accurately convert data from 2D intersections into 3D volumetric estimates, it is necessary to assume something about the nature of crystal shapes (Higgins, 2006). To determine the best shape approximation, we described data from slabs and thin sections using common shape parameters: edge ‘roughness’ (P/A), circularity ($4\pi \cdot \text{area} / (P^2)$), angularity (P/(P of equivalent area circle)), and oblateness ($1 - \text{major axis} / \text{minor axis}$) (Fig. 7). The smallest size fraction from each scale of observation are not included in this analysis, due to the effect of pixilation on these shape parameters: for thin sections, $ol > 0.04$ mm; for slabs, $ol > 0.1$ mm. These truncated datasets indicate that the vast majority of olivine crystals are circular to prolate ellipsoidal

in shape, do not have irregular or concave edges, and have aspect ratios between 1.5:1 and 1:1 (Fig. 7). Olivine crystals show variability in circularity with respect to size: olivine crystal data collected at the thin section scale (0.01 – 10 mm in diameter) show distinctly more crystals with angular and less circular geometries than crystal data collected at the slab scale (0.8– 10 mm) (Fig. 7b). Image processing resolution (11.93 pixels/mm) may have a pronounced effect on circularity values, captured in part by the ‘clustering’ of circularity vs. size data for the smallest olivine crystals (Fig. 7b). However, 2D intersections may not give an accurate representation of the true crystal aspect ratios (3D); better approximations can be obtained using the CSDSliceTM database developed by Morgan and Jerram (2006). In Morgan and Jerram’s database, a minimum of 10,000 slices are taken through poly-disperse (*sensu* Higgins, 2000), randomly-oriented model populations created from over 700 crystal shapes; these data are compiled in a database for comparison with 2D intersection data obtained in real rocks. This comparison showed the best-fit, 3D, long : int : short aspect ratios to be ~1.5:1:1 for data collected at the thin section scale, and 1:1:1 for random samples of 4000 crystals from normalized and truncated olivine crystal populations (Tables A1, A2; Table 1). We further test the raw data for the effect of crystal size on the best-estimate for true crystal aspect ratio (Fig. 7d): though slightly variable, best-approximations of true 3D aspect ratios for all size bins do not exceed 1.5:1:1. Correlation coefficients (R^2) of best-fit approximations for 3D aspect ratios are very low (<0.8), and thus these 3D aspect ratios are not considered representative (Tables A1, A2; Table 3; Fig.6d). However, combined with observations of olivine crystal morphologies, these results suggest that, although no single shape is representative of olivine crystal morphologies, smaller olivines are less circular and can have more irregular external morphologies, and most olivine crystals have an equant geometry. For our stereological conversions, therefore, we elect to use the best-approximation long : int : short ratio of 1:1:1.

Applications of spherical conversion coefficients to non-spherical particles always results in over-estimation of the number of ‘small’ particles and under-estimation of ‘large’ particles. However, using spherical geometry assumptions instead of assuming various object shapes with unequal axes lengths also restricts possible errors of estimating 3D populations from 2D surfaces (Sahagian and Proussevitch, 1998).

Binning. Olivine crystal size data are sorted into two bin types by diameter: \log_2 (corresponding to integer ϕ values), and $\log_{0.5}$ (corresponding to typical diamond sieve sizes), each capturing the entire observed population of olivine crystals (0.03 – 12 mm or -3 to $+5\phi$). This allows us to examine normalized, truncated olivine crystal size data in a classic histogram format (Figs. 6a,b), to perform stereological conversions of 2D (area) data to 3D (volume) (Table 2), and to subsequently examine the effect of bin size on representations of olivine crystal size data (Fig.8a).

Section-effect corrections. To account for the intersection-probability and cut-section effects, we stereologically convert binned, 2D crystal data (maximum diameter of enclosing ellipse) into 3D using a shape assumption in CSDCorrectionsTM. We assume all olivine crystals are equant and approximately spherical in three dimensions with 3D aspect ratios of 1:1:1, and apply the following correction to the logarithmically-binned major axes of olivine crystals:

$$n_V(L_{X-Y}) = \frac{n_A(l_{X-Y})}{D} \quad (\text{after Royet, 1991}) \quad (\text{A1})$$

where $n_V(L_{X-Y})$ is the number of crystals per unit volume for a given length scale, $n_A(l_{X-Y})$ is the number of crystals per unit area for a given length scale, and D is the minimum diameter of the spheres in a given size interval. This correction is only applied to bins with diameter > 0.03 mm; for the few crystals smaller than this size, observed crystal outlines are likely projections of the whole crystal outline, as these diameters are thinner than a typical thin-section. Using the number densities of crystals per unit volume obtained by stereological correction above $[n_V(L_{X-Y})]$, we find the population density by use of the following equation:

$$N(L_{X-Y}) = \frac{n_V(L_{X-Y})}{W_{X-Y}} \quad (\text{after Higgins, 2006}) \quad (\text{A2})$$

where the population density $N(L_{X-Y})$ is the number density of crystals per unit volume in the size interval, $n_V(L_{X-Y})$, divided by the width of the size bin, W_{X-Y} . To take into account the cut-section effect, we use the variation of the Saltikov method (Saltikov, 1967) in CSDCorrectionsTM 1.3 for unfolding a population of intersection lengths into their true lengths by using a function of the intersection lengths (after Higgins, 2000; Higgins, 2006; Sahagian and Proussevitch, 1998). Further corrections involving volumetric phase abundance (based on modal area %), crystal roundness, and measurement style (major enclosing ellipse) are employed in CSDCorrectionsTM 1.3 to generate cumulative crystal size distributions.

The resulting logarithmically-binned and stereologically-corrected data described above are used to construct several types of CSD's of olivine in coherent kimberlite. The majority of CSD's published in petrology literature use 'classic' semi-logarithmic 'CSD diagrams', in which population density is defined by size intervals. We plot the CSDs from the five samples of coherent kimberlite in a semi-logarithmic classic CSD to allow for comparison with other crystal populations from igneous rocks (Fig. 8a), and bi-logarithmic $\ln N$ vs. \ln size (Fig. 8b) to assess an appropriate model distribution.

Population verification

To verify our calculation of the 3-D CSD from 2-D data, we use two independent measures of the modal abundance (as %) of olivine crystals: (1) point-counting in thin sections and (2) integration of the crystal volumes resulting from our CSD calculations. By comparing these two data types with the initial measured area of crystals (e.g., Table 1), we can verify if the scale integration and CSD corrections applied to our initial data set are in compliance with the measurements initially made by the image analysis program (i.e., area %). Point-counting is based on the understanding that the fraction of random or equally spaced points that lie on a phase, P_p is also equal to the volumetric fraction of the phase, V_V , and thus represents a stereologically exact 'global parameter' (Higgins, 2006). One thousand evenly-spaced points on two thin sections were point-

counted, and the resulting data are shown in Table A3. The volumetric proportion of olivine, V_v , is calculated by integration of the volume of all the crystals in all bin sizes of our olivine CSD using the following:

$$V_v = \sigma \int_0^{\infty} n_v(L) L^3 dL \quad (\text{after Higgins, 2006}) \quad (\text{A3})$$

where σ is the ratio of the crystal volume (assumed spheres) to that of a cube of side L . Higher order moments need to use shape factors (σ) to account for departures of the crystal shapes from a cube (Higgins, 2006). As our data are modeled as spheres with diameters equivalent to the area measured, we use $\pi/6$ for σ . The relative volumetric proportions of olivine (V_{x-y}) of different size bins are similar to the relative area % obtained by normalizing initial observations made in thin sections and slabs to the slab scale (Table A3). These results are also similar to the data obtained by point-counting from a thin section taken from the same slab (Table A3), suggesting that observations of modal abundance made at the thin section scale approximate larger scales of observation.

Goodness-of-fit tests:

To evaluate the hypothesis that each CSD fits a power-law distribution model, we first identify the best-fit power law equation and subsequently evaluate how well binned data from each sample population correspond to the model equations. This evaluation is accomplished by calculating a weighted least-squares fit to the sample data and root mean standard error (RMSE) for best-fit lines to the data (Davis, 1986).

Data are first plotted as $\log N$ vs. $\log L$, where N_i is the population density in the i th bin size, and L is the diameter (mm) of the maximum axis. These linear trends are fitted using a weighted, linear least squares fit on data with the following modification of equation (1) to find the slope (D) and intercept (λ) of a straight line:

$$\log N_v = -D \log x + \log \lambda \quad (\text{A4})$$

Errors on each $\log x$, $\log y$ data point range from 0.17 to 0.29. We compute slopes (D) ± 10 , and intercepts (λ) ± 10 , and plot the covariance ellipse at the 95% confidence level (Figure 8c). This shows the range of model slopes and intercepts (D, λ) that are fully consistent (at 95% confidence) with the data and their uncertainty. The variance on the model parameters for samples within a single dyke (e.g., $\sigma_{CK4,5p}^2$) are smaller than variance between dykes ($\sigma_{pCK1,2,3; pCK4,5}^2$). The variance in parameters for all samples (σ_{sp}^2) is greater than the variance attributed to analytical methods (σ_{ap}^2), such that $\sigma_{ap}^2 < \sigma_{sp}^2$.

We also computed the root mean standard error (RMSE) as a measure of the differences (i.e. residuals) between values expected by the best-fit power-law models and values observed from the olivine crystal populations. The RMSE is a good measure of

accuracy, as it measures the variability expected in the means of samples by repeated random collection from the same population:

$$RMSE = \sqrt{\sum_{i=1}^n \left(\frac{(O_{\lambda} - (M_D \times O_D) - M_{\lambda})^2}{n} \right)} \quad (A5)$$

where n is the number of bins, M_D and M_{λ} the expected model parameter values, and O_{λ} and O_D the observed parameter values from the iterations. RMSE error estimates for power-law fits to the CSDs are shown in Table 3. An RMSE close to zero means the model is a good predictor of the observed data; RMSE for all 5 samples in this study range from 0.16 to 0.31, which accords well with the estimated average uncertainties ($\sigma_{tr} + \sigma_{thn}$) on each data point (0.17-0.29).