Supplementary Information for: **Metrics of graininess: robust quantification of grain-count from the non-uniformity of scattering rings**

July 21, 2014

Kevin G. Yager (kyager@bnl.gov), Pawel W. Majewski
Center for Functional Nanomaterials, Brookhaven National Laboratory, Upton, NY 11973 U.S.A

1 Relative Deviation of a Single Peak

We consider the average and standard deviation for a single Gaussian peak appearing in the signal $I(\chi)$. Without loss of generality, we assume the angular linecut is along $[-\pi, +\pi)$ with a single peak centered at $\chi = 0$:

$$I_p(\chi) = I(0)e^{-\chi^2/2\sigma_\chi^2} \tag{1}$$

The average is:

$$\mu_p = \langle I_p(\chi) \rangle = \frac{\int_{-\pi}^{+\pi} I(0)e^{-\chi^2/2\sigma_\chi^2} \mathrm{d}\chi}{\int_{-\pi}^{+\pi} \mathrm{d}\chi} \tag{2}$$

$$= \frac{I(0)\sigma_\chi\sqrt{2\pi}}{2\pi} \tag{3}$$

$$= \frac{I(0)\sigma_\chi}{\sqrt{2\pi}} \tag{4}$$

Note that we have made the reasonable assumption that the peak is well-contained within the integration range ($3\sigma_\chi < \pi$), so that the full Gaussian integral is a good approximation for the

finite integration. The variance is then:

$$\begin{aligned}
\sigma_p^2 &= \frac{\int_{-\pi}^{+\pi} \left(I_p(\chi) - \mu_p\right)^2 \mathrm{d}\chi}{\int_{-\pi}^{+\pi} \mathrm{d}\chi} && (5)\\
&= \frac{1}{2\pi} \int_{-\pi}^{+\pi} \left(I(0) e^{-\chi^2/2\sigma_\chi^2} - \frac{I(0)\sigma_\chi}{\sqrt{2\pi}} \right)^2 \mathrm{d}\chi && (6)\\
&= \frac{I(0)^2}{2\pi} \int_{-\pi}^{+\pi} \left(e^{-2\chi^2/2\sigma_\chi^2} - \frac{2\sigma_\chi}{\sqrt{2\pi}} e^{-\chi^2/2\sigma_\chi^2} + \frac{\sigma_\chi^2}{2\pi} \right) \mathrm{d}\chi && (7)\\
&= \frac{I(0)^2}{2\pi} \left[\frac{\sigma_\chi}{\sqrt{2}} \sqrt{2\pi} - \frac{2\sigma_\chi}{\sqrt{2\pi}} \sigma_\chi \sqrt{2\pi} + \frac{\sigma_\chi^2}{2\pi} 2\pi \right] && (8)\\
&= \frac{I(0)^2 \sigma_\chi}{2\pi} \left[\sqrt{\pi} - 2\sigma_\chi + \sigma_\chi \right] && (9)\\
\sigma_p &= \frac{I(0)\sigma_\chi}{\sqrt{2\pi}} \sqrt{\frac{\sqrt{\pi}}{\sigma_\chi} - 1} && (10)
\end{aligned}$$

2 Relative Deviation of Multiple Peaks

A realistic signal $I(\chi)$ of course includes many peaks. We can analyze the relative deviation of this case also. The intensity is given by:

$$I(\chi) = \sum_{i=1}^{N_p} I(0) e^{-(\chi-\chi_i)^2/2\sigma_\chi^2} \tag{11}$$

The average is simply:

$$\begin{aligned}
\mu_I = \langle I(\chi) \rangle &= \frac{\int_{-\pi}^{+\pi} \sum_{i=1}^{N_p} I(0) e^{-(\chi-\chi_i)^2/2\sigma_\chi^2} \mathrm{d}\chi}{\int_{-\pi}^{+\pi} \mathrm{d}\chi} && (12)\\
&= \frac{I(0)}{2\pi} \sum_{i=1}^{N_p} \int_{-\pi}^{+\pi} e^{-(\chi-\chi_i)^2/2\sigma_\chi^2} \mathrm{d}\chi && (13)\\
&= \frac{I(0) N_p \sigma_\chi \sqrt{2\pi}}{2\pi} && (14)\\
&= \frac{I(0) N_p \sigma_\chi}{\sqrt{2\pi}} && (15)
\end{aligned}$$

The variance is then:

$$\sigma_I^2 = \frac{\int_{-\pi}^{+\pi} (I(\chi) - \mu_I)^2 \, \mathrm{d}\chi}{\int_{-\pi}^{+\pi} \mathrm{d}\chi} \tag{16}$$

$$= \frac{1}{2\pi} \int_{-\pi}^{+\pi} I(\chi)^2 \mathrm{d}\chi - \mu_I^2 \tag{17}$$

$$= \frac{I(0)^2}{2\pi} \int_{-\pi}^{+\pi} \left(\sum_{i=1}^{N_p} e^{-(\chi-\chi_i)^2/2\sigma_\chi^2} \right)^2 \mathrm{d}\chi - \mu_I^2 \tag{18}$$

$$= \frac{I(0)^2}{2\pi} \int_{-\pi}^{+\pi} \left(\sum_{i=1}^{N_p} e^{-2(\chi-\chi_i)^2/2\sigma_\chi^2} + \sum_{i \neq j} e^{-(\chi-\chi_i)^2/2\sigma_\chi^2} e^{-(\chi-\chi_j)^2/2\sigma_\chi^2} \right) \mathrm{d}\chi - \mu_I^2 \tag{19}$$

$$= \frac{I(0)^2}{2\pi} \left(N_p \sigma_\chi \sqrt{\pi} + \sum_{i \neq j} \int_{-\pi}^{+\pi} e^{-(\chi-\chi_i)^2/2\sigma_\chi^2} e^{-(\chi-\chi_j)^2/2\sigma_\chi^2} \mathrm{d}\chi \right) - \mu_I^2 \tag{20}$$

The remaining integrand can be written as:

$$\exp\left(-(\chi-\chi_i)^2/2\sigma_\chi^2 - (\chi-\chi_j)^2/2\sigma_\chi^2\right) \tag{21}$$

$$= \exp\left(-\frac{1}{2\sigma_\chi^2} \left((\chi-\chi_i)^2 + (\chi-\chi_j)^2 \right) \right) \tag{22}$$

$$= \exp\left(-\frac{1}{2\sigma_\chi^2} \left(\chi^2 - 2\chi\chi_i + \chi_i^2 + \chi^2 - 2\chi\chi_j + \chi_j^2 \right) \right) \tag{23}$$

$$= \exp\left(-\left(\frac{1}{\sigma_\chi^2} \right) \chi^2 \right) \exp\left(-2 \left(-\frac{\chi_i + \chi_j}{2\sigma_\chi^2} \right) \chi \right) \exp\left(-\frac{\chi_i^2 + \chi_j^2}{2\sigma_\chi^2} \right) \tag{24}$$

Note that $\int_{-\infty}^{\infty} e^{-ax^2} e^{-2bx} \, \mathrm{d}x = \sqrt{\frac{\pi}{a}} e^{\frac{b^2}{a}}$, so:

$$\sigma_I^2 = \frac{I(0)^2}{2\pi} \left(N_p \sigma_\chi \sqrt{\pi} + \sum_{i \neq j} \sqrt{\pi \sigma_\chi^2} e^{(\chi_i+\chi_j)^2/4\sigma_\chi^2} e^{(\chi_i^2+\chi_j^2)/2\sigma_\chi^2} \right) - \mu_I^2 \tag{25}$$

$$= \frac{I(0)^2 \sigma_\chi}{2\sqrt{\pi}} \left(N_p + \sum_{i \neq j} e^{(\chi_i+\chi_j)^2/4\sigma_\chi^2} e^{(\chi_i^2+\chi_j^2)/2\sigma_\chi^2} \right) - \mu_I^2 \tag{26}$$

The remaining summation includes $N_p^2 - N_p$ terms. If N_p is small, then the peaks will tend to not overlap, and so the summation of cross-peak terms will be negligible. In this case we obtain:

$$\sigma_I^2 = \frac{I(0)^2 \sigma_\chi}{2\sqrt{\pi}} N_p - \mu_I^2 \tag{27}$$

$$= \frac{I(0)^2 \sigma_\chi}{2\sqrt{\pi}} N_p - \frac{I(0)^2 \sigma_\chi^2}{2\pi} N_p^2 \tag{28}$$

$$= \frac{I(0)^2 \sigma_\chi^2 N_p^2}{2\pi} \left(\frac{\sqrt{\pi}}{\sigma_\chi N_p} - 1 \right) \tag{29}$$

Which yields:

$$\sigma_R = \frac{\sigma_I}{\mu_I} = \sqrt{\frac{\sqrt{\pi}}{\sigma_\chi N_p} - 1} \tag{30}$$

Since N_p is small, this scaling matches the expected $N_p^{-1/2}$. In the more general case, the summation includes terms that depend on the peak positions, which in turn depend on the orientation distribution. For a random orientation distribution, we expect the $N_p^{-1/2}$ scaling to be maintained, as observed in the simulation results.

3 Grain size distribution

3.1 Lognormal distribution

The number-averaged distribution of grain sizes for a lognormal distribution can be written:

$$\mathrm{d}N = \frac{1}{\xi\omega\sqrt{2\pi}} e^{-(\ln\xi-\gamma)^2/2\omega^2}\mathrm{d}\xi \tag{31}$$

Where γ and ω are the mean and standard deviation of the logarithm of the distribution. The number-average of the distribution is:

$$\langle\xi\rangle_N = e^{\gamma+\frac{\omega^2}{2}} \tag{32}$$

The standard deviation is:

$$\Delta_{\xi_N} = \langle\xi\rangle_N\sqrt{e^{\omega^2}-1} \tag{33}$$

The area-weighted mean is:

$$\langle\xi\rangle_A = e^{\gamma+\frac{5\omega^2}{2}} \tag{34}$$

And the volume-weighted mean is:

$$\langle\xi\rangle_V = e^{\gamma+\frac{7\omega^2}{2}} \tag{35}$$

We can write:

$$\omega = \sqrt{\ln\left(\frac{\langle\xi\rangle_N^2}{\Delta_{\xi_N}^2}+1\right)} \tag{36}$$

$$\gamma = \ln\langle\xi\rangle_N - \frac{\omega^2}{2} \tag{37}$$

$$= \ln\langle\xi\rangle_N - \frac{1}{2}\ln\left(\frac{\langle\xi\rangle_N^2}{\Delta_\xi^2}+1\right) \tag{38}$$

Since $\langle\xi\rangle_V = \langle\xi\rangle_N e^{3\omega^2}$, we can also write:

$$\frac{\Delta_{\xi_N}}{\langle\xi\rangle_V} = \frac{\sqrt{e^{\omega^2}-1}}{e^{3\omega^2}} \tag{39}$$

$$\gamma = \ln\frac{\langle\xi\rangle_V}{e^{3\omega^2}} - \frac{\omega^2}{2} \tag{40}$$

We use the above relations to construct lognormal distributions that have matched volume-weighted average by using the parameter ω to vary the width of the distribution, and then computing γ in order to maintain $\langle\xi\rangle_V$ (Fig. S1).

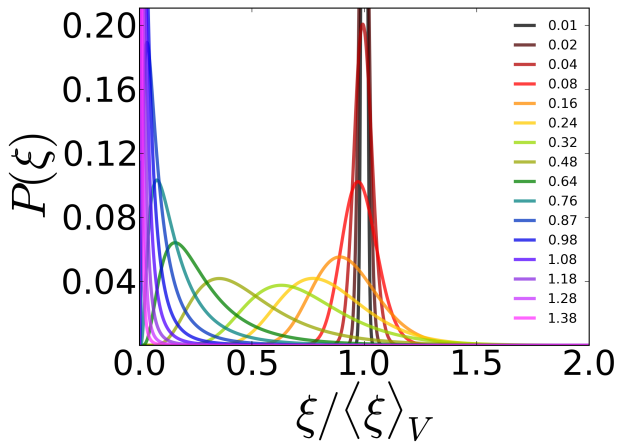

Figure S1: The lognormal distribution, shown for a variety of distribution widths (modulated by the parameter ω, as indicated in the legend). The parameter γ is adjusted so as to keep the volume-weighted average ($\langle\xi\rangle_V$) constant. For $\omega < 0.3$, the distribution is pseudo-Gaussian. For larger ω, the distribution becomes increasingly skewed.

3.2 Scaling of σ_R due to grain size

Recall that the metric σ_R follows the functional form:

$$\sigma_R = c_{\sigma_R} N_g^{\beta_{\sigma_R}} \tag{41}$$

The prefactor can be expressed in terms of σ_χ, or in terms of the dimensionless grain size $q\langle\xi\rangle_V$ (which is essentially the number of lattice repeats in the grain, up to a factor 2π) by using the Scherrer relation $\langle\xi\rangle_V = 2\pi K/\sigma_q$, which can also be written $\sigma_\chi = 2\pi K/q\langle\xi\rangle_V$:

$$\begin{aligned}
c_{\sigma_R}^2 &= \frac{2(\sqrt{\pi} - \sigma_\chi)}{m\sqrt{2\pi}\sigma_\chi^2 f_\chi} && (42)\\
&= \frac{2(\sqrt{\pi} - \frac{2\pi K}{q\langle\xi\rangle_V})}{m\sqrt{2\pi}\left(\frac{2\pi K}{q\langle\xi\rangle_V}\right)^2 f_\chi} && (43)\\
&= \frac{2(\sqrt{\pi} q\langle\xi\rangle_V - 2\pi K) q\langle\xi\rangle_V}{m\sqrt{2\pi}\,(2\pi K)^2 f_\chi} && (44)\\
&= \frac{(q\langle\xi\rangle_V - 2\sqrt{\pi}K) q\langle\xi\rangle_V}{m2\sqrt{2}\pi^2 K^2 f_\chi} && (45)\\
&= \frac{1}{m2\sqrt{2}\pi^2 K^2 f_\chi}\left(q\langle\xi\rangle_V\right)^2 - \frac{1}{m\sqrt{2}\pi^{3/2} K f_\chi} q\langle\xi\rangle_V && (46)
\end{aligned}$$

Note that since the first term is considerably larger than the second, this relation can be approximated by the linear relation:

$$c_{\sigma_R} = \frac{q\langle\xi\rangle_V}{K\pi\sqrt{2\sqrt{2}m f_\chi}} \tag{47}$$

The full expression can also be solved via the quadratic equation to give:

$$
\begin{aligned}
q\langle\xi\rangle_V &= +\frac{m2\sqrt{2}\pi^2K^2f_\chi}{2m\sqrt{2}\pi^{3/2}Kf_\chi} \pm \frac{m2\sqrt{2}\pi^2K^2f_\chi}{2}\sqrt{\left(\frac{1}{m\sqrt{2}\pi^{3/2}Kf_\chi}\right)^2 - 4\left(\frac{1}{m2\sqrt{2}\pi^2K^2f_\chi}\right)\left(-c_{\sigma_R}^2\right)} \\
&= K\sqrt{\pi} \pm \frac{m\sqrt{2}\pi^2K^2f_\chi}{1}\sqrt{\frac{1}{m^2 2\pi^3K^2f_\chi^2} + \frac{4c_{\sigma_R}^2}{m2\sqrt{2}\pi^2K^2f_\chi}} \qquad (48) \\
&= K\sqrt{\pi} \pm m\pi^2K^2f_\chi\sqrt{\left(\frac{2}{2mf_\chi\pi^2K^2}\right)\left[\frac{1}{m\pi f_\chi} + \frac{4c_{\sigma_R}^2}{\sqrt{2}}\right]} \qquad (49) \\
&= K\sqrt{\pi} \pm K\pi\sqrt{\left(\frac{mf_\chi}{1}\right)\left[\frac{1}{m\pi f_\chi} + \frac{2\sqrt{2}c_{\sigma_R}^2}{1}\right]} \qquad (50) \\
&= K\sqrt{\pi} + K\sqrt{\pi + 2\pi^2\sqrt{2}mf_\chi c_{\sigma_R}^2} \qquad (51)
\end{aligned}
$$

Where we have taken the positive root as the physically meaningful one. Moreover, this expression can be approximated by the linear expression:

$$
q\langle\xi\rangle_V = K\pi\sqrt{2\sqrt{2}mf_\chi}c_{\sigma_R} \qquad (52)
$$

3.3 Phenomenological trends for σ_R due to grain size distribution

In cases where there exists a distribution in grain size, the $\sigma_R = c_{\sigma_R}N_g^{\beta_{\sigma_R}}$ becomes modified; in particular we note that β_{σ_R} deviates from $-1/2$. Simulations were performed where the scattering peak intensities and widths are both affected by grain size: integrated peak intensities were scaled with grain volume, and peak width was modulated using the Scherrer relation. Figure S2 shows the results of one such simulation. The scaling of the metrics is modified, owing to the additional variation which the grain-size distribution introduces. In particular, the power-law exponent (slope in the log-log graph) of the σ_R scaling becomes smaller ($|\beta_{\sigma_R}| < 1/2$). In a given scattering image, the larger grains will contribute preferentially to the observed intensity fluctuation; owing both to their larger scattering intensity (due to the larger grain volume), and their higher peak intensity (owing to their peak sharpness). Overall, as additional grains are added to the scattering volume, the variation in the signal is thus not averaged as rapidly (because of the additional variation of the grain-size distribution).

We generated a wide range of simulations, varying the grain size and distribution. We assess the behavior of c_{σ_R} and β_{σ_R} by phenomenologically fitting the variation of these parameters with (dimensionless) grain size $q\langle\xi\rangle_V$ and breadth of the size distribution (quantified by the lognormal parameter ω). Since we are not enforcing $\beta_{\sigma_R} = -1/2$, we observe a correspondingly modification to the behavior of c_{σ_R}. In particular, for sufficiently broad size distributions, the observed scaling becomes shallower and c_{σ_R} deviates substantially from the theoretical behavior already derived (Fig S3).

The observed trends can be described using phenomenological equations of the form:

$$
\begin{aligned}
\beta_{\sigma_R} &= -\frac{1}{2}\left(1 - \tanh\left[\left(\frac{\omega}{\delta_\beta}\right)^{\kappa_\beta}\right]\right) \qquad (53) \\
c_{\sigma_R} &= \frac{q\langle\xi\rangle_V}{K\pi\sqrt{2\sqrt{2}mf_\chi}}\left(1 - \tanh\left[\left(\frac{\omega}{\delta_c}\right)^{\kappa_c}\right]\right) \qquad (54)
\end{aligned}
$$

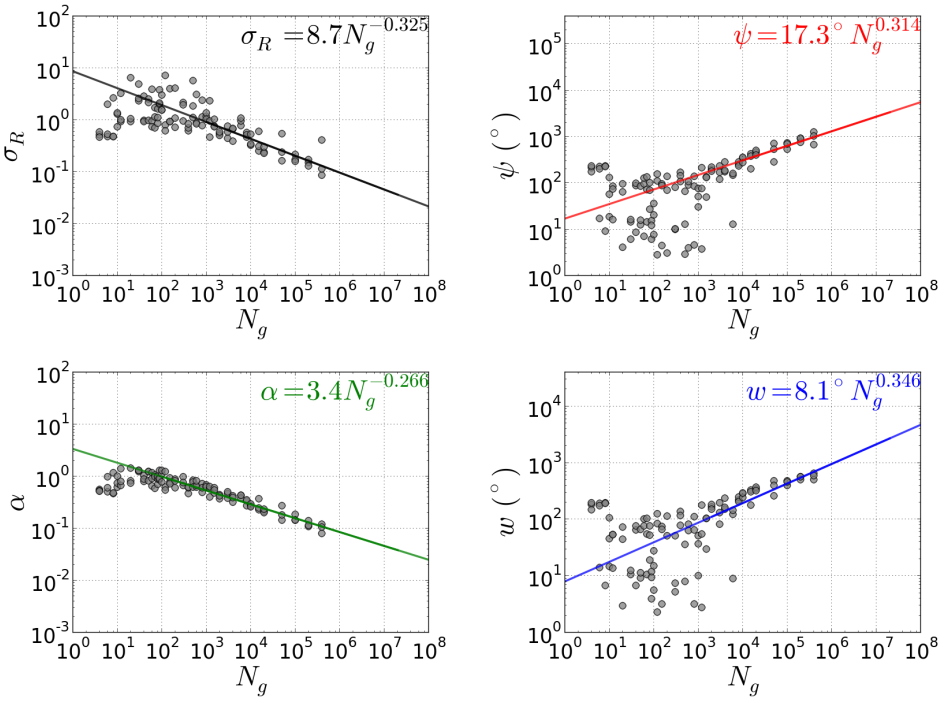

Figure S2: Scaling behavior of the metrics as a function of the number of grains, N_g, for a system with a lognormal distribution of grain sizes (simulations shown are for $q\langle\xi\rangle_V = 261$, $\omega = 0.64$ and $B = 0.001$). The power-law scaling is modified, compared to the case where there is no grain-size distribution; in particular the exponent of the σ_R scaling becomes shallower ($|\beta_{\sigma_R}| < 1/2$).

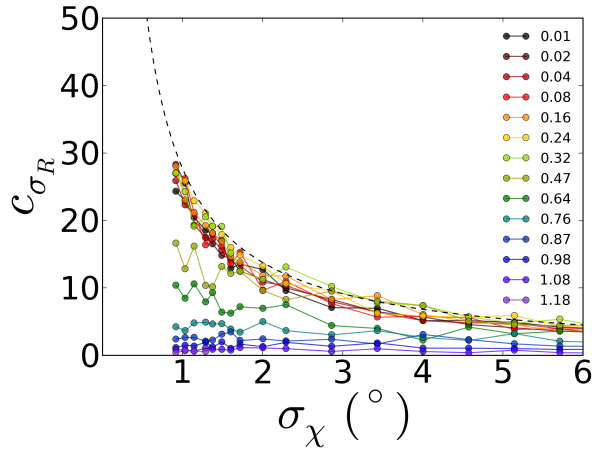

Figure S3: Variation of the σ_R power-law prefactor for different widths of the grain size distribution (legend denotes ω). In this data, the grain size was allowed to affect both the peak heights and peak widths (through the Scherrer relation) in the scattering data (i.e. the grains are sufficiently small that peaks widths are not dominated by instrumental resolution).

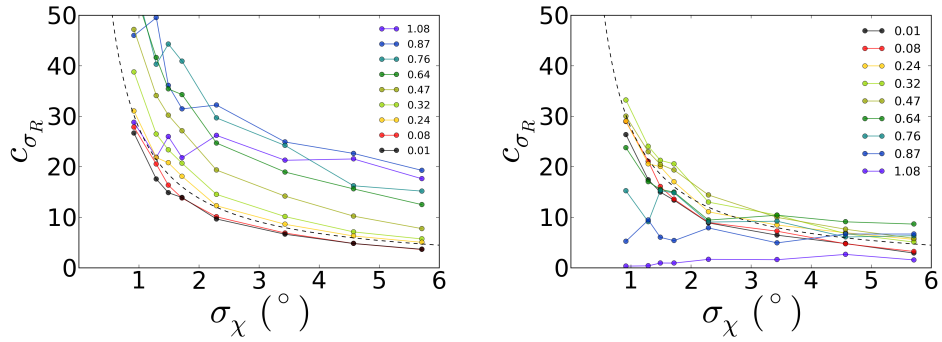

Figure S4: Variation of the σ_R power-law prefactor for different widths of the grain size distribution (legend denotes ω), in the case where the grain sizes are sufficiently large that they no longer affect the observed peak width. The left panel is the values of c_{σ_R} obtained when fitting the simulation result while enforcing $\beta = -0.5$, whereas the right panel shows the values obtained when allowing β to vary freely in the fit.

where tanh is the hyperbolic tangent function. These equations can be used to fit the simulation data, where we observe that the phenomenological fit parameters are roughly $\delta \approx 1$ and $\kappa \approx 2$. In fact, the entire set of simulation curves can be fit in a concerted fashion, which yields $\delta_\beta = 1.02$, $\kappa_\beta = 2.32$, $\delta_c = 0.83$, and $\kappa_c = 2.30$. This single set of parameters adequately describes the full set of 221 simulations performed (see main text). However, the fit is purely phenomenological, and we observe additional, more subtle, trends in the data. E.g. β_{σ_R} appears to have a weak positive dependence on $q\langle\xi\rangle_V$, and the c_{σ_R} similarly have some additional weak dependence on ω, not accounted for in the master curves presented here. In principle, this higher-order influence could be accounted for by allowing the phenomenological parameters to vary between conditions; i.e. allow $\delta_\beta(q\langle\xi\rangle_V)$ and so on. Nevertheless, this simple phenomenological model is sufficient to use the simulation results as a calibration curve, for the interpretation of experimental data. We note also that the deviations from ideal behavior only become relevant for very broad grain-size distributions.

3.4 Phenomenological trends for σ_R for fixed peak widths

The above results were obtained when both peak intensities and widths were affected by grain size. In cases where the average grain size is large, the intrinsic peak width will be much smaller than the instrumental peak width. We now therefore consider a set of simulations wherein only the peak heights (and not the peak widths) are influenced by the grain size (i.e. valid in the limit where instrumental broadening always dominates). We again observe the power-law prefactor varies from its idealized value (Fig. S4). However, by allowing β to vary during the fit, we can capture the portion of the signal variability arising from the distribution of grain-sizes. We observe that the same phenomenological relations already described can be used to fit these trends. However, the κ exponents are considerably larger, which in fact makes the analysis procedure less sensitive to grain-size distribution when the distribution is relatively narrow. In other words, the proposed analysis should be robust for realistic grain-size distributions, when the grains are so large that the peak widths are instrumental-dominated. This comparative insensitivity likely arises from the suppression of one additional source of variation (namely, variation of the peak width). On the other hand, we note that the deviation from the proposed phenomenology appears to be more extreme in

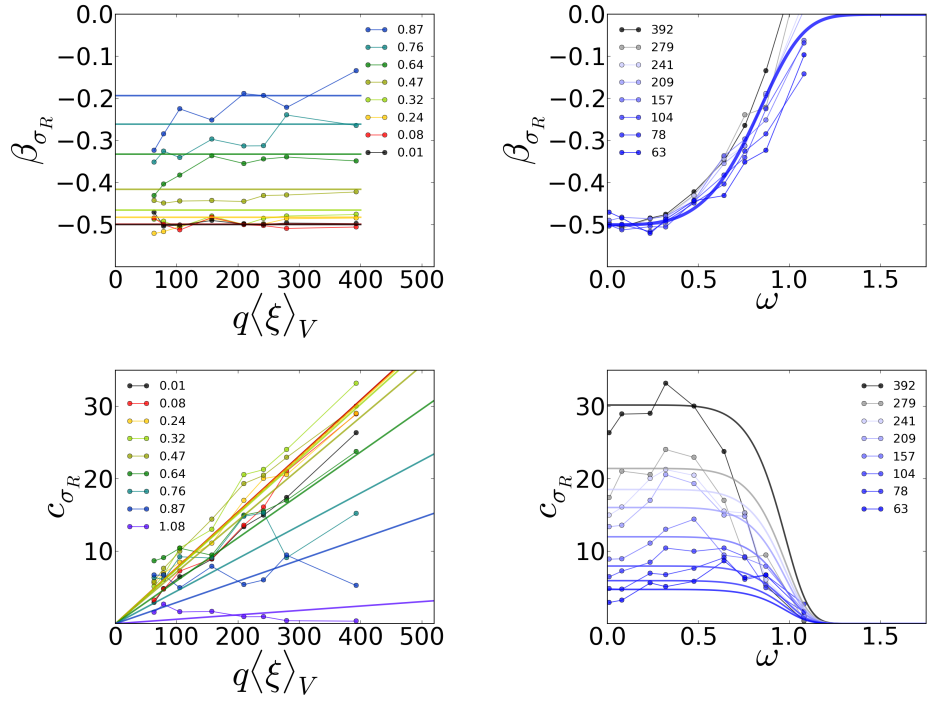

Figure S5: Variation of the σ_R power-law fitting variables. The left panels show the variation as a function of grain sizes ($q\langle\xi\rangle_V$), for different grain size distributions (ω given in legend). The right panels show the same data plotted versus grain size distribution ($q\langle\xi\rangle_V$ given in legend). The solid lines are concerted phenomenological fits to the entire simulation dataset, which yields $\delta_\beta = 0.95$, $\kappa_\beta = 3.65$, $\delta_c = 1.03$, and $\kappa_c = 6.72$..

this case of fixed peak width. With respect to using the simulation results as a calibration curve, one can simply perform a phenomenological fit on the relevant subset of the data.

3.5 Scaling of ψ due to grain size

The preceding analysis can of course be revisited for the other metrics. Similarly to the analysis for σ_R, we know that the metric ψ follows: $\psi = c_\psi N_g^{\beta_\psi}$. Converting our relation for the prefactor to be in terms of the dimensionless grain size:

$$
\begin{aligned}
c_\psi &= 4\sigma_\chi^{3/2}\sqrt{\frac{m}{2}\sqrt{2\pi}f_\chi} && (55)\\
&= 4\left(\frac{2\pi K}{q\langle\xi\rangle_V}\right)^{3/2}\sqrt{\frac{m}{2}\sqrt{2\pi}f_\chi} && (56)\\
&= 8\pi^{3/2}\frac{K^{3/2}}{\left(q\langle\xi\rangle_V\right)^{3/2}}\sqrt{m\sqrt{2\pi}f_\chi} && (57)\\
q\langle\xi\rangle_V &= \left(8\pi^{3/2}\frac{K^{3/2}}{c_\psi}\sqrt{m\sqrt{2\pi}f_\chi}\right)^{2/3} && (58)\\
&= 4\pi\frac{K}{c_\psi^{2/3}}\left(m\sqrt{2\pi}f_\chi\right)^{1/3} && (59)
\end{aligned}
$$