On the scaling of multicrystal data sets collected at high-intensity X-ray and electron sources

Philip Coppens and Bertrand Fournier
Chemistry Department, University at Buffalo, State University of New York, Buffalo, New York 14260-3000, USA

(Received 10 September 2015; accepted 30 October 2015; published online 11 November 2015)

The need for data-scaling has become increasingly evident as time-resolved pump-probe photocystallography is rapidly developing at high intensity X-ray sources. Several aspects of the scaling of data sets collected at synchrotrons, XFELs (X-ray Free Electron Lasers) and high-intensity pulsed electron sources are discussed. They include laser-ON/laser-OFF data scaling, inter- and intra-data set scaling.

I. INTRODUCTION

With the increasing application of pump-probe photocystallography over a range from milliseconds to femtoseconds, sample crystals have increasingly needed to be replaced due to laser-induced radiation damage. Even more dramatic, X-ray induced damage is dominant in femtosecond experimentation as performed at X-ray free electron laser (XFEL) sources and high brightness electron diffraction facilities, leading the sample to explode immediately after the generation of a diffraction pattern. In that case, the subsequent analysis must be based on multi-crystal data sets, which implies that data-scaling becomes a crucial aspect of the processing of the data, as has been recognized in many reported studies, see, for example, Hattne et al. (2014) and Kirian et al. (2011). We will discuss three types of scaling and provide references to more explicit descriptions where necessary.

II. CONCEPTS

Time-resolved pump-probe experiments can be performed by repeated measurements on a reversible system with different pump-probe delays if necessary, or, in serial time-resolved crystallography, by following non-reversible processes, such as chemical or biological reactions at different time points. In the first case, the RATIO method can be used in which laser-ON and laser-OFF frames in a single setting are collected in rapid succession, and in the second case, the time-zero dark and serial light-on structures have to be collected separately.

Important quantities in time-resolved crystallography are the response ratio $R$, defined for each reflection $H$ as

$$R(H) = \frac{I^{\text{ON}}(H)}{I^{\text{OFF}}(H)},$$

where $I^{\text{ON}}(H)$ and $I^{\text{OFF}}(H)$ are, respectively, the laser-ON and laser-OFF reflection intensities, and the relative response $\eta(H)$, defined as
\[ \eta(H) = \frac{I_{\text{ON}}(H) - I_{\text{OFF}}(H)}{I_{\text{OFF}}(H)} = R(H) - 1. \]  

(1)

In the RATIO method of data analysis, the ON/OFF ratios \( R \) are the basic quantities on which the refinement is based (Coppens et al., 2009).

One of the aims of the studies is the calculation of photodifference maps, showing the light-induced change in electron density, either as a function of time if a dynamic process is being recorded or as due to light induced electronic excitation (Zheng et al., 2007; Makal et al., 2011; and Fournier and Coppens, 2014b). Such maps, defined as

\[ \Delta \rho(r) = \rho^{\text{ON}}(r) - \rho^{\text{OFF}}(r) = \sum_{H \in \text{Hunter et al.}} [|F^{\text{ON}}(H)| - |F^{\text{OFF}}(H)|] e^{-2\pi i H \cdot r} e^{i \phi^{\text{OFF}}(H)}, \]  

(2a)

require ON and OFF structure factors. Experimental values may be used for \( |F^{\text{ON}}(H)| \) and \( |F^{\text{OFF}}(H)| \), or alternatively \( |F^{\text{OFF}}(H)| \) may be obtained from a structure model refined against an accurate independently collected reference data set, and the \( |F^{\text{ON}}| \) data derived from multiplication by the square root of the experimentally determined ratios. In that case

\[ |F^{\text{ON}}(H)| = |F^{\text{OFF}}(H)| \sqrt{R(H)} = |F^{\text{OFF}}(H)| \sqrt{1 + \eta(H)}, \]  

(3)

which gives for expression (2)

\[ \Delta \rho(r) = \sum_{H} |F^{\text{OFF}}(H)| \sqrt{R(H)} - |F^{\text{OFF}}(H)| e^{-2\pi i H \cdot r} e^{i \phi^{\text{OFF}}(H)} \]  

\[ = \sum_{H} |F^{\text{OFF}}(H)| \sqrt{R(H)} - 1 | e^{-2\pi i H \cdot r} e^{i \phi^{\text{OFF}}(H)}. \]  

(2b)

The use of structure factors from an accurately refined reference model minimizes the experimental errors in (2b) which will then essentially result from the ratio determination. The maps include the effect of structural rearrangements and temperature changes induced by the laser exposure, the former including electron density changes on excitation, the latter however typically being a much smaller effect than that of the features due to the atomic motions. An example of a photodifference map, showing density changes resulting from molecular excitation of a \([\text{CuI}((1,10\text{-phenanthroline-N,N}) \text{bis(triphenylphosphine)}) \text{ complex (Makal et al., 2012)}, \) is shown in Figure 1. A comparable map on a dirhodium complex can be found in a 2011 publication (Makal et al., 2011).

FIG. 1. 90K photodifference maps with isosurfaces of ±0.25 e Å\(^{-3}\) (red positive, blue negative). The crystal has two distinct molecules in the asymmetric unit, which distort differently. Reproduced with permission from Makal et al., J. Phys. Chem. A 116, 3359–3365 (2012). Copyright 2012 International Union of Crystallography.
III. TYPES OF SCALING

A. ON/OFF scaling

For each frame pair \( I \) collected with the consecutive ON/OFF exposure technique with different incident beam intensities, a scale factor can be defined as

\[
K^I = \frac{I_{\text{beamON}}}{I_{\text{beamOFF}}},
\]

(4)

The scaled ON/OFF ratio for each \( H \) in the frame pair \( I \) is obtained from

\[
R_{\text{scaled}}(H) = \frac{R(H)}{K^I} = \frac{I^{\text{ON}}(H)\sigma_{\text{beamON}}}{I^{\text{OFF}}(H)\sigma_{\text{beamOFF}}},
\]

(5)

For ON/OFF pairs of well populated frames, this scale factor can be estimated by analysis of the very low-order reflections, as the \( \eta \)'s will go to zero in the forward scattering direction, in which all electrons scatter in phase. For example, the linear regression y-intersect of the photo-Wilson plot (Schmøkel et al., 2010) can be used to estimate the factor \( K^I \). This argument neglects any increase in diffuse scattering on exposure, which may be significant if conversion percentages are large. Once the difference between the incident beam intensities of the laser-ON and laser-OFF frames has been corrected, the inter-set scaling can be performed using the procedures described in Section III B.

1. ON/OFF scaling, special cases

In the time-resolved synchrotron experiment, it is feasible to collect laser-ON and laser-OFF frames in rapid succession, thus making sure that crystal quality and other conditions, such as incident beam intensity or crystal size, are the same. This eliminates the need for relative scaling of the ON and OFF frames, which is otherwise necessary for calculation of the ratios in expression 2b, following expression (5). This is not the case when radiation damage becomes excessive at XFEL sources. When a liquid jet injector is used to pass a stream of nanocrystals through the photon beam (Spence et al., 2012), no consecutive ON/OFF exposure is feasible as the crystals are destroyed immediately after the diffracted beams have been emitted. However, the nanocrystals may be stationarily positioned in arrays, which may be done by use of a crystallization chip with multiple basins (Mueller, 2014), an ultra-thin silicon nitride membrane (Hunter et al., 2014), or similar. In this case, it is possible to take a first exposure with attenuated incident beam intensity below the destruction limit to collect the OFF-frame, to be followed by a second exposure with full intensity for the ON-frame (Mueller et al., 2015). The consecutive ON/OFF-exposure technique is advantageous even in this case as the OFF and ON reflections are measured on the same sample crystal, but the method still requires the ON/OFF scaling described above, as the OFF beam has been attenuated to preserve the sample and apart from that, the X-ray pulses at XFELs vary in intensity. If the ON and OFF data have been collected on different samples, data set merging and scaling must be performed first separately on all ON and all OFF data after each of the patterns has been successfully indexed, prior to ON/OFF scaling.

B. Inter-set scaling

Two different scaling methods have been developed. Both are based on the assumption that the effects of the thermal increase and that of the induced structural changes on excitation are proportional to the laser intensity and penetration depth. The Absolute-Average-System-Response (AASR) Method is simple and fast. It has been briefly described earlier and applied to synchrotron data sets on a bridged binuclear Rh compound (Makal et al., 2011). The more
sophisticated \textit{Weighted Least-Squares (WLS) Method} is based on the fit of a non-structural model to the intensity changes \( \eta \) in the different data sets.

\textbf{1. The AASR method}

The \( \eta \) values can be both positive and negative. But if the effect of the thermal increase and that of the induced structural changes on excitation are proportional, the average absolute value of the \( \eta \)s over each of the data set should be scaled to a common value to put the \( \eta \) values in different sets on a common scale. Thus, a scale factor can be defined for each set \( i \) as

\[ Z_i = \frac{\langle \eta_{\text{obs}}(H) \rangle_{\text{all sets}}}{\langle \eta_{\text{obs}}(H) \rangle_{\text{set } i}} \]  

(6)

to give

\[ \eta_{\text{obs}}^{\text{scaled}}(H) = Z_i \eta_{\text{obs}}^i(H). \]  

(7)

The method is fast, but it relies apart from the assumption of proportionality of structural and thermal effects, on a considerable overlap of the reflections in the various data sets, especially in cases in which the structural changes, and therefore, the scattering changes, are very anisotropic.

\textbf{2. The WLS method}

The weighted least-squares method as applied to the \( \eta \) values (Fournier and Coppens, 2014a), defined in expression (1), follows a procedure described by Hamilton in 1965 (Hamilton \textit{et al.}, 1965) for the layer scaling of Weissenberg photographs and further developed by Fox and Holmes (Fox and Holmes, 1966).

The function minimized is defined as

\[ \varepsilon_{\text{min}}^2(x) = \sum_{i \in \{ \text{sets} \}} \left[ \sum_{H \in \{H\}} \left( \sum_{j \in \{\eta_{\text{obs}}\}} w_{\text{obs}}^{(i,j)}(H) (\eta_{\text{obs}}^{(i,j)}(H) - \langle \eta_{\text{model}}^i(H) \rangle)^2 \right) \right], \]  

(8)

in which the parameter-based model for \( \eta(H) \) in set \( i \), \( \eta_{\text{model}}^i(H) \) is given by \( Q_i \langle \eta_{\text{model}}^i \rangle(H) \), \( \langle \eta_{\text{model}} \rangle(H) \) being the calculated average \( \eta \) for reflection \( H \) over all sets, \( w_{\text{obs}}^{(i,j)} \) is the weight of reflection \( j \) in set \( i \), and \( Q_i \), the relative excited-state population, is defined as \( Q_i = P_i / \langle P \rangle_{\text{all sets}} \), with \( P \) the excited state population in set \( i \).

\( \langle \eta_{\text{model}} \rangle(H) \) and \( Q_i \) are the variables in the WLS refinement. The scaled \( \eta \) values of set \( i \) are then obtained from

\[ \eta_{\text{obs}}^{\text{scaled}}^i(H) = \frac{\eta_{\text{obs}}^i(H)}{Q_i}. \]

The values of \( \langle \eta_{\text{model}} \rangle(H) \) can be used in the calculation of the photodifference map defined in expressions (2a) and (2b).

\textbf{3. Intra-data set scaling}

A logical consequence of the anisotropic optical nature of all but cubic crystals is the anisotropy of light absorption and therefore of light-induced changes such as photoreactions or molecular excitation. In addition, when samples with dimensions of several tens of microns or larger are used in the experiments, light penetration is typically limited, which will cause an orientation dependence of the light absorption in irregularly shaped samples.

In a molecular crystal, the optical anisotropy is in first approximation dominated by the molecular properties. The main absorption direction represented by the transition dielectric dipole moment integral vector is a common output of theoretical calculations. As a sample is
rotated during data collection or when different crystals are used for collection of single frames, the angle between the transition moment and the incident beam will change, leading to vari- 
ation of the absorption. To minimize (but not eliminate) the effect, the laser pump-beam should be circularly rather than linearly polarized, as illustrated in Figure 2.

The intensity of absorption at an instant $t$ is proportional to $|E|^2|\mu|^2\cos^2\theta$ (Malus law), in which $E$ is the vector of the electric component of the laser beam. Integration over the electric field plane of the circularly polarized laser leads to an average intensity of the absorption proportional to $\sin^2\gamma$. Clearly, the effect is strongly dependent on the relative orientation of crystal. When the sample will be rotated around the laser direction, no variation due to absorption anisotropy should be observed. On the other hand, when the sample is rotated perpendicular to the laser direction, the anisotropy may be pronounced.

IV. EXAMPLE OF AN APPLICATION

A synchrotron study was done on the tetranuclear complex $\text{Ag}_2\text{Cu}_2\text{L}_4$ ($\text{L} = 2 – \text{diphenylphosphino} – 3\text{–methylindole}$) (Jarzembska et al., 2014) at beamline 14-ID of BioCars at APS. The complex shows significant structural changes on excitation by 355 nm laser light, the silver-silver contact in the complex shortening by 0.38(3) Å on ligand-to-metal charge transfer. Four data sets were collected on different samples. Details are summarized in Table I.

Merging of the four data sets leads to a global completeness of 46.7% with a redundancy of multiply measured and equivalent ON/OFF ratios of 5.4. The achieved resolution is controlled by instrumental limitations to a maximal $\sin \theta/\lambda$ value of 0.61 Å$^{-1}$.

Results of inter- and intra- data set scaling of the four sets of data collected are illustrated in Figure 3. The horizontal red and black lines represent the effect of the inter-data set scaling.

| Set | Power (mJ/mm$^2$) | Strategy angle range | Usable angle range | Number obs. ratios | Number unique hkl | Complet. (%) | Redundancy | Max. Reso. ($\sin \theta/\lambda$) |
|-----|------------------|----------------------|-------------------|--------------------|------------------|-------------|------------|---------------------|
| 1$^a$ | 0.50 | 0–90 | 0–18;20–69 | 9403 | 3158 | 36.8 | 3.0 | 0.526 Å$^{-1}$ |
| 2 | 0.40 | 0–90 | All | 5243 | 2887 | 24.2 | 1.8 | 0.587 |
| 3 | 0.25 | 0–180 | All | 12 708 | 5295 | 39.2 | 2.4 | 0.611 |
| 4 | 0.25 | 0–90 | All | 6744 | 3621 | 29.8 | 1.9 | 0.591 |

$^a$Unique data set collected with a pair redundancy $\langle N \rangle = 10$ and $\phi$ angle steps $\Delta \phi = 1^\circ$, while all others collected with $\langle N \rangle = 5$ and $\Delta \phi = 2^\circ$. 

**TABLE I. Description of the $\text{Ag}_2\text{Cu}_2\text{L}_4$ data sets.**
with the WLS and AASR methods, respectively. For data set 3 both horizontal lines superimpose. Inter-set scaling differences are apparent especially for sets 1 and 2. However, the effect of intra-data set scaling is much more pronounced especially for set 3 as shown by the curved lines in black and red, respectively, obtained using the AASR and WLS methods. In the case of the WLS method, a restraint is added to the WLS error function to reduce the variation between adjacent $\phi$ settings, which smoothes the lines.

The inter-scaling relative populations obtained with the different scaling methods described above are summarized in Table II.

The atomic shifts on excitation obtained in a joint refinement against the four individual sets are summarized in Table III. The results are remarkably independent of the details of the analysis.

The robustness of the results is likely due to the large redundancy averaging at 5.4 in the experimental ratios. In a joint refinement against the four data sets, as described above, the effect of equivalent reflections being affected differently will average out in the final results.

### A. Intra-data set scaling of set 3

The pronounced variation of the experimental ratios with the $\phi$-angle allows an experimental determination of the direction of the transition moment integral vector $\mu$. Application of
**TABLE III.** Atomic shifts in the refined structure models obtained by joint model refinements performed with the program LASER2010 (Vorontsov et al., 2010) against independently merged data sets with and without intra-set scaling using the scaling methods AASR or WLS. The atomic shifts reported in the original paper (Jarzembska et al., 2014) are listed in the first column.

| Set  | Reported in paper | No scaling | AASR method | WLS method |
|------|-------------------|------------|-------------|------------|
| Ag1  | 0.30(2)           | 0.31(2)    | 0.31(2)     | 0.32(2)    |
| Ag2  | 0.27(1)           | 0.27(2)    | 0.27(2)     | 0.27(2)    |
| Cu1  | 0.09(2)           | 0.13(3)    | 0.12(3)     | 0.12(2)    |
| Cu2  | 0.33(2)           | 0.34(3)    | 0.35(3)     | 0.36(2)    |

**TABLE IV.** Theoretical (HSEH1PBE)/LANL2DZ results of the transition moment integral directions for a molecule of Ag$_2$Cu$_2$L$_4$ placed in the crystallographic unit cell. TD-DFT (Time-Dependent Density Functional Theory) results using the Gaussian (PBE1PBE)/LANL2DZ method are essentially identical.

| Transitions   | Transition dipole moment direction | Energy (eV) | Oscillator strength |
|---------------|-----------------------------------|-------------|--------------------|
| HOMO → LUMO   | −0.961                            | 0.217       | 0.169              | 2.5124 | 0.0156 |
| HOMO-1 → LUMO | −0.928                            | −0.303      | 0.219              | 2.5834 | 0.0258 |
| HOMO-2 → LUMO | 0.112                             | 0.640       | 0.760              | 2.7062 | 0.0149 |
| HOMO-3 → LUMO | −0.592                            | 0.359       | 0.721              | 2.8051 | 0.0039 |

$\sin^2\gamma$-dependence to the WLS data of set 3, discussed in Section III B 3 leads to a direction equal to

$$\mu_{\text{normalized}} = [-0.955 ; 0.073 ; -0.287].$$

This may be compared with the following values from Gaussian/DFT(HSEH1PBE)/LANL2DZ calculations for the four lowest energy transitions (Table IV). The first two transitions, which have the highest oscillator strengths, have a transition-moment direction oriented close to the a-axis, in agreement with the experimental result. The agreement between theory and experiment is quite reasonable and supports the validity of the experimental methods used.

**V. CONCLUSIONS**

We conclude that the anisotropy of absorption followed by molecular excitation can be pronounced, but as expected, depends strongly on the orientation of the sample with respect to the laser beam. When a large number of individual data sets are collected as necessitated by sample destruction in a very intense X-ray beam, and to a much lesser extent by laser damage in synchrotron experiments, anisotropy effects, evident from intra-data set scaling, may average in joint refinements, or in a unique-set refinement performed when all sets are merged. Inter-data set scaling is essential if photodifference maps are to be examined. ON/OFF scaling must be used when the consecutive ON/OFF-exposure method, applied to the same sample with the same incident beam intensity, cannot be used.

**ACKNOWLEDGMENTS**

Support of this work by the National Science Foundation (CHE-1213223) was gratefully acknowledged. Use of the BioCARS Sector 14 was supported by the National Institutes of Health, National Center for Research Resources, under Grant No. RR007707. The Advanced Photon Source was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, under Contract No. W-31-109-ENG-38.
Coppens, P., Pitak, M., Gembicky, M., Messerschmidt, M., Scheins, S., Benedict, J., Adachi, S.-I., Sato, T., Nozawa, S., Ichiyanagi, K., Chollet, M., and Koshihara, S.-Y., J. Synchrotron Radiat. 16, 226–230 (2009).

Fournier, B. and Coppens, P., Acta Crystallogr., Sect. A: Found. Adv. 70, 514–517 (2014a).

Fournier, B. and Coppens, P., Acta Crystallogr., Sect. A: Found. Adv. 70, C771 (2014b).

Fox, G. C. and Holmes, K. C., Acta Crystallogr. 20, 886–891 (1966).

Hamilton, W. C., Rollett, J. S., and Sparks, R. A., Acta Crystallogr. 18, 129–130 (1965).

Hattem, J., Echols, N., Tran, R., Kern, J., Gildea, R. J., Brewster, A. S., Alonso-Mori, R., Glockner, C., Hellmich, J., Laksmono, H., Sierra, R. O., Lassalle-Kaiser, B., Lampe, A., Han, G., Gul, S., DiFiore, D., Milathianaki, D., Fry, A. R., Miahnabri, A., White, W. E., Schafer, D. W., Seibert, M. M., Koglin, J. E., Sokaras, D., Weng, T. C., Sellberg, J., Lutimers, M. J., Glatzel, P., Zwart, P. H., Grosse-Kunstleve, R. W., Bogan, M. J., Messerschmidt, M., Williams, G. J., Boutet, S., Messinger, J., Zouni, A., Yano, J., Bergmann, U., Yachandra, V. K., Adams, P. D., and Sauter, N. K., Nat. Methods 11, 545–548 (2014).

Hunter, M. S., Segelke, B., Messerschmidt, M., Williams, G. J., Zatsepin, N. A., Barty, A., Benner, W. H., Carlson, D. B., Coleman, M., Graf, A., Hau-Riege, S. P., Pardini, T., Seibert, M. M., Evans, J., Boutet, S., and Frank, M., Sci. Rep. 4, 5 (2014).

Jarzembska, K., Kaminski, R., Fournier, B., Trzep, E., Sokolow, J., Henning, R., Chen, Y., and Coppens, P., Inorg. Chem. 53, 10594–10601 (2014).

Kirian, R. A., White, T. A., Holton, J. M., Chapman, H. N., Fromme, P., Barty, A., Lomb, L., Aquila, A., Maia, F. R. N. C., Martin, A. V., Fromme, R., Wang, X., Hunter, M. S., Schmidt, K. E., and Spence, J. C. H., Acta Crystallogr. 67, 131–140 (2011).

Makal, A., Benedict, J., Trzep, E., Sokolow, J., Fournier, B., Chen, Y., Kalinowski, J. A., Graber, T., Henning, R., and Coppens, P., J. Phys. Chem. A 116, 3359–3365 (2012).

Makal, A., Trzep, E., Sokolow, J. D., Kalinowski, J., Benedict, J. B., and Coppens, P., Acta Crystallogr., Sect. A: Found. Crystallogr. 67, 319–326 (2011).

Mueller, C., Ph.D. thesis, University of Toronto (Canada), ProQuest Dissertations Publishing, 2014.

Mueller, C., Marx, A., Epp, S. W., Zhong, Y., Kuo, A., Balo, A. R., Soman, J., Schotte, F., Lemke, H. T., Owen, R. L., Pai, E. P., Pearson, A. R., Olson, J. S., Anfinrud, P. A., Ernst, O. P., and Dwayne Miller, R. J., Struct. Dyn. 2, 054302 (2015).

Schmokel, M., Kaminski, R., Benedict, J. B., and Coppens, P., Acta Crystallogr., Sect. A: Found. Crystallogr. 66, 632–636 (2010).

Spence, J. C. H., Weierstall, U., and Chapman, H. N., Rep. Prog. Phys. 75, 102601 (2012).

Vorontsov, I., Pillet, S., Kaminski, R., Schmokel, M., and Coppens, P., J. Appl. Cryst. 43, 1129–1130 (2010).

Zheng, S.-L., Messerschmidt, M., and Coppens, P., Acta Crystallogr., Sect. B: Struct. Sci. 63, 644–649 (2007).