A generalized method for high throughput in-situ experiment data analysis:

An example of battery materials exploration

Bachir Aoun¹, Cun Yu¹, Longlong Fan¹, Zonghai Chen², Khalil Amine², Yang Ren¹

¹Advanced Photon Source, Argonne National Laboratory, Argonne, IL 60439

²Chemical Science and Engineering Division, Argonne National Laboratory, Argonne, IL 60439

Abstract

A generalized method is introduced to extract critical information from series of ranked correlated data. The method is generally applicable to all types of spectra evolving as a function of any arbitrary parameter. This approach is based on correlation functions and statistical scedasticity formalism. Numerous challenges in analyzing high throughput experimental data can be tackled using the herein proposed method. We applied this method to understand the reactivity pathway and formation mechanism of a Li-ion battery cathode material during high temperature synthesis using in-situ high-energy X-ray diffraction. We demonstrate that Pearson's correlation function can easily unravel all major phase transition and, more importantly, the minor structural changes which cannot be revealed by conventionally inspecting the series of diffraction patterns. Furthermore, a two-dimensional (2D) reactivity pattern calculated as the scedasticity along all measured reciprocal space of all successive diffraction pattern pairs unveils clearly the structural evolution path and the active areas of interest during the synthesis. The

methods described here can be readily used for on-the-fly data analysis during various in-situ operando experiments in order to quickly evaluate and optimize experimental conditions, as well as for post data analysis and large data mining where considerable amount of data hinders the feasibility of the investigation through point-by-point inspection.

Introduction

The development of faster sensors, detectors and instruments has been continuously accelerating data acquisition speed and enlarging data volume in the field of both basic and applied research. The availability of state-of-the-art high throughput tools and techniques, especially at large user facilities like Advanced Photon Source (APS), has significantly facilitated the in-situ operando exploration of advanced materials in real time and realistic conditions. However, due to the rapid data collection and large data volume, the conventional approach to analyze and inspect each spectrum becomes inefficient and sometimes unfeasible. It is particularly important for researchers to quickly evaluate results during the experiment, in order to decide if there is any critical parameter or region for improved experimental conditions and better results. There is an urgent and increasing demand for a fast and universal method to interrogate high throughput experimental data sets for both on-the-fly and post data analysis. Such a method will help uncover data insights to accelerate scientific discovery and innovation. In this paper, we report a generalized method based on correlation functions and statistical scedasticity formalism to extract important information of series of spectra as a function of any type of variable. We demonstrate its capability by applying this method to explore the reactivity pathway and

formation mechanism during high temperature synthesis, which are widely used for making cathode materials for Li-ion batteries.

Lithium transition metal oxides are dominant cathode materials for lithium-ion batteries that power the modern portable electronics. Substantial amount of effort has been devoted to formulating an optimal chemical composition and a proper process to maximize the electrochemical performance and life of such cathode materials [1, 2, 3, 4, 5, 6, 7]. Most of the incremental improvement reported in the open literature is mostly accomplished by try-and-error approach [4, 5, 6, 7] due to the lack of a systematic and fundamental study of material behavior during the preparation of such materials. This has been a common dilemma in the general field of the material science. The high-intensity, high-quality X-ray source generated at cutting-edge synchrotron facilities has enabled us to investigate the real materials in real devices at a realistic condition. The latest development of temporal/special resolved techniques provide us unprecedented multidimensional characterization power to establish the structure-property relationship of functional materials [8, 9, 10]. However, the full utilization of such instrumental power is barely achieved because the prior knowledge/experience is generally sought after for manual data reduction before a feasible in-depth analysis can be carried out. This sometime leads to a systematic “ignorance” of some subtle information that is of great scientific value.

Recently, *in situ* synchrotron probes are extensively utilized to investigate batteries materials during the synthesis [11, 12, 13], during the normal operation of batteries [14, 15], as well as during the abuse of battery materials [16, 17, 15]. The latest development of synchrotron technologies enable us to collect thousands of spectra for a single experiment, and the exhaustive processing of the experimental data becomes less and less possible. Very often, one needs to identify the most important region of interest for detailed data analysis [17, 15]. Alternatively,

computer-based mathematical tools are often pursued for the data reduction and data process to minimize the unnecessary influence of prior knowledge. For instance, Cui et. al. carried out *in situ* X-ray Absorption Near Edge Spectroscopy (XANES) to study the (de)lithiation mechanism of SeS_x anode material for lithium-ion batteries. Using the residual error of linear combination analysis as an indicator, Cui et. al. was able to identify the existence and the chemical composition of intermediate compounds (Li_2Se_2 and Li_2Se_4) during the charge/discharge of SeS_x anode [18]. Similarly, Piao et. al. used factor analysis (FA) to process the *in situ* XANES data during the charge/discharge of LiVPO_4F cathode, and was able to predict the existence of an intermediate phase $\text{Li}_x\text{VPO}_4\text{F}$ during the charge/discharge process [19]. Kan et. al. used evolutionary factor analysis (EFA) to analyze the *in situ* XANES data during the solid state synthesis of Li_2MnO_3 using MnCO_3 and Li_2CO_3 as the starting materials. Without the need of prior knowledge, the EFA was able to extract the evolution profile and the pure XANES spectrum of an intermediate compound, determined to be MnO_2 , which was not able to be characterized by *in situ* high energy X-ray diffraction (HEXRD) [20]. In this work, we utilize *in situ* HEXRD to track the structural evolution during the solid state synthesis of $\text{LiNi}_{0.7}\text{Mn}_{0.15}\text{Co}_{0.15}\text{O}_2$, a promising high energy-density cathode material for lithium-ion batteries. Due to the complexity and non-linear nature of XRD data, above linear-algebra-based algorithms are not suitable for processing *in situ* HEXRD data. Therefore, Pearson's correlation functions and statistical scedasticity formalism are proposed herein to mathematically detect the subtle change of HEXRD patterns during the *in situ* experiment.

Experimental details

The powders of $(\text{Mn}_{0.15}\text{Ni}_{0.7}\text{Co}_{0.15})\text{CO}_3$ and Li_2MnO_3 were mixed with a molar ratio of 2:1, and cool-pressed into a pellet, about 2 mm in thickness and 7 mm in diameter. The precursor $(\text{Mn}_{0.15}\text{Ni}_{0.7}\text{Co}_{0.15})\text{CO}_3$ was prepared using a co-precipitation method reported in our previous papers [1, 2]. The pellet was then placed in a Linkam high temperature furnace TS1500. The in-situ high-energy X-ray diffraction experiment with a transmission mode was performed at the beam line 11-ID-C of the Advanced Photon Source (APS), at Argonne National Laboratory. The wavelength of the X-rays was 0.11165 Å. Two-dimensional (2D) HEXRD patterns were collected continuously using a Perkin-Elmer large area detector when the sample was heated up from room temperature up to 850 °C with a constant heating rate of 1 °C per minute. The 2D patterns were calibrated and converted to the conventional patterns of intensity versus 2-theta using Fit2d software. The temperature dependent HEXRD patterns are shown as a contour plot in Fig. 1. Details of the experimental setup and data treatment can be found in our previous report [17].

Correlation analysis

Correlation methods are commonly used to extract signal from background noise and to measure dependence between two sets of data. Noda has discussed the procedure of generating 2D correlation spectra from infrared spectroscopy [21, 22, 23] and he has shown that a two-dimensional covariance map can be calculated for any ranked spectral data sets (equation 18 in ref 23) [23]. Faguy and Balachander [24] showed that by analyzing infrared spectroscopy data using Noda's two. dimensional covariance analyses, it becomes possible to determine methyl and methylene stretching modes of lithium perchlorate diethyl ether electrolyte solutions where it

was previously thought to show no solute solvent dependent bands in the mid-Infrared spectra. Here our aim is to unravel the fine details and to easily identify the interesting temperature dependent in-situ X-rays diffraction data ranges that need more attention and studying. Firstly, we calculated Noda's normalized covariance matrix ' CM ' (equation 1.1).

$$CM = [CM_{i,j}]_{i,j=1 \rightarrow N}^{1 \rightarrow 1 \rightarrow N} \quad (1.1), \quad CM_{i,j} = \frac{\int \vec{I}(2\theta, t_i) * \vec{I}(2\theta, t_j) d\theta}{\sqrt{\int \vec{I}^2(2\theta, t_i) d\theta * \int \vec{I}^2(2\theta, t_j) d\theta}} \quad (1.2)$$

Where ' N ' is the number of temperatures at which a diffraction pattern has been collected, ' i ' and ' j ' give the position in the covariance matrix corresponding respectively to temperatures ' t_i ' and ' t_j ', ' 2θ ' is the diffraction angle, ' $\vec{I}(2\theta, t)$ ' is the diffraction pattern collected at constant temperature t and ' T ' is the transpose symbol. ' $CM_{i,j}$ ' (equation 1.2) is the matrix covariance element between diffraction patterns at temperatures ' t_i ' and ' t_j ', its value reflects how changes in the diffraction pattern at temperature ' t_i ' correlate with changes at temperature ' t_j '. The temperature dependent covariance map for the HEXRD patterns is plotted in Fig. 2. It's important to note that because of the commutative property of the dot product, the covariance matrix is symmetric with respect to its main diagonal ' $CM_{i,j} = CM_{j,i}$ '. Each element of the main diagonal is the auto-correlation of the diffraction pattern ' $\vec{I}(2\theta, t_i)$ ' at temperature ' t_i ' its value is therefore equal to 1 ' $CM(i,i) = 1$ '

As a result one should look at off diagonal lines ' $CM(i, i+l)_{i,l \in N}$ ' because they indicate correlated changes at ' $t_{i+l/2} \Big|_{l=1}^{l=N-l}$ ' along the temperature range ' $[t_i, t_{i+l}]_{l=1}^{l=N-l}$ '. The first off diagonal line is shown in Fig. 3 corresponding to ' $l = 1$ '. Higher order diagonals ' $l > 1$ ' compute

the covariance between ' l ' separated diffraction patterns. In general, the likelihood of strong de-correlation increases by increasing the order ' l ' and significant subtle changes recorded between successive patterns can be completely missed. Therefore for high quality data, it's advised to analyze ' $l = 1$ ' covariance because it provides finer details and reflects meaningful and important structural changes.

Changes in the 2D diffraction patterns (Fig. 1) are well reflected in the covariance map (Fig. 2) and the off diagonal lines (Fig. 3). At 260°C, the transition metal carbonate, $(\text{Mn}_{0.15}\text{Ni}_{0.7}\text{Co}_{0.15})\text{CO}_3$, decomposed to transition metal oxides with the NiO structure, as revealed by a series of new peaks at 2θ (deg.) = 2.65, 3.06, 4.33, 5.08, 5.31, ... corresponding to reflection (111), (200), (220), (311), (222), ... of a rock salt structure with the symmetry of $Fm-3m$, while the peaks of MeCO_3 disappear. From about 380°C, the transition metal oxides start to react with Li_2CO_3 to form LiMeO_2 (Me= $\text{Mn}_{0.15}\text{Ni}_{0.7}\text{Co}_{0.15}$) type layered structure, which is completed at about 770°C with Li_2CO_3 totally consumed. In order to track down the microstructure evolution of the layered cathode materials, we performed detailed structural analysis of the diffraction patterns between 450°C and 825°C using the Rietveld method. At 450°C, there are three phases coexisting, so we use the following phases to the refinement: Li_2CO_3 with the symmetry $C2/c$, MeO with $Fm-3m$ symmetry and $(\text{Li}_{1-x}\text{Ni}_x)(\text{Me}_{1-x}\text{Li}_x)\text{O}_2$ with $R-3m$. For the layered cathode material, we assumed that there is ion exchange between the two crystallographic sites for Li and transition metal. The observed and calculated patterns at 450°C with the three-phase fitting are plotted in Fig. 4a. The temperature dependence of the weight fraction of each phase is plotted in Fig. 4b, which shows that Li_2CO_3 and MeO phases are gradually diminishing with increasing temperature. Fig. 4c shows the temperature dependent variation of the lattice parameters of the layered cathode materials and MeO. Interestingly, we

found that the temperature dependence of the lattice parameters of the $(\text{Li}_{1-x}\text{Ni}_x)(\text{Me}_{1-x}\text{Li}_x)\text{O}_2$ phase has a non-monotonic variation during material growth. Although X-rays are not very sensitive to the light elements, like Li, in comparison with other heavy elements, we plot the Li concentration vs. temperature in Fig. 4d. Again, it shows an additional feature during the synthesis. Such a feature is not captured by the above mentioned covariance approach.

In order to be able to investigate more details of the structural evolution, we need to find a more sensitive and direct method to nab subtle changes in large series of spectra. Here we propose to use Pearson's correlation coefficient to calculate the dependence between XRD spectra at consecutive temperatures. While both the covariance map and Pearson's correlation function reveal whether two sets of data are positively or inversely correlated, Pearson's correlation coefficient is more sensitive and provides additional information because it is normalized by the standard deviation therefore it reflects the degree of linear relationship between two spectra. It may take any value between -1 and 1. A correlation value of 0 indicates that there is no association between the two spectra at all; positive correlation value means that the two spectra follow the same trend, while negative correlation means that when one spectrum is increasing the second one is more likely to decrease and vice-versa. Pearson's correlation designed by the letter ' ρ ' is defined as the covariance of the two spectra divided by the product of their standard deviations equation (2)

$$\rho_{t,t+\tau} = \frac{\text{cov}(\vec{I}(2\theta, t), \vec{I}(2\theta, t + \tau))}{\sigma_t \sigma_{t+\tau}} = \frac{\int (\vec{I}(2\theta, t) - \mu_\theta(t)) * (\vec{I}(2\theta, t + \tau) - \mu_\theta(t + \tau)) d\theta}{\sqrt{\int (\vec{I}(2\theta, t) - \mu_\theta(t))^2 * (\vec{I}(2\theta, t + \tau) - \mu_\theta(t + \tau))^2 d\theta}}$$

$$\rho_{t_i, t_{i+1}} = \frac{\sum_{i=1}^N (\tilde{I}(2\theta, t_i) - \mu_\theta(t_i)) * (\tilde{I}(2\theta, t_{i+1}) - \mu_\theta(t_{i+1}))}{\sqrt{\sum_{i=1}^N (\tilde{I}(2\theta, t_i) - \mu_\theta(t_i))^2 * (\tilde{I}(2\theta, t_{i+1}) - \mu_\theta(t_{i+1}))^2}} \quad (2)$$

Where ' $\mu_\theta(t)$ ' is the expectancy value of ' $\tilde{I}(2\theta, t)$ ', in our case, we merely calculated it as the mean value of the scattering pattern at temperature ' t ' (equation 3).

$$\mu_\theta(t) = \langle \tilde{I}(2\theta, t) \rangle = \frac{1}{\theta_{\max} - \theta_{\min}} \int_{\theta_{\min}}^{\theta_{\max}} \tilde{I}(2\theta, t) d\theta = \frac{1}{N_\theta} \sum_{i=1}^{N_\theta} \tilde{I}(2\theta_i, t) \quad (3)$$

In addition to first off diagonal features in the covariance matrix, the Pearson's correlation function calculated between all successive spectra shows other structural de-correlations at different temperatures (Fig. 3). Furthermore, in the inset of Fig. 3, we show a zoom-in view of the region between 400°C and 800°C. Pearson's function exhibits a de-correlation peak at around 560°C that appears to be smeared out in the covariance first off diagonal. A superposed plot (Fig. 5) of the two successive diffraction spectra at 548°C and 550°C seems to be not enough to reveal the structural reason behind the loss of correlation at 550°C. Therefore, we introduce another new statistical method in order to investigate the homogeneity of variance between two spectra along the ' 2θ ' axis, we define the scedasticity function (equation 4) as the partitioned correlation between two signals.

$$s_{t_i, t_{i+1}}(2\theta) = \frac{1}{|\tilde{I}(2\theta, t_i)| * |\tilde{I}(2\theta, t_{i+1})|} \int_{2\theta - \frac{\Delta(2\theta)}{2}}^{2\theta + \frac{\Delta(2\theta)}{2}} \tilde{I}(2\theta, t_i) * \tilde{I}(2\theta, t_{i+1}) d\theta \quad (4)$$

Where ' $\Delta(2\theta)$ ' is the partition window size in 2θ space. The scedasticity function between the two diffraction spectra at 548°C and 550°C calculated with a window $\Delta(2\theta) = 0.05^\circ$ clearly shows the activity range as a de-correlation peak around 2.25 degrees (Fig. 6). By inspecting this region near the strong de-correlation peak (Fig. 6), we found that it was related to the decomposition of the Li_2CO_3 phase at a temperature above 400°C (Fig. 6, inset). We have calculated the scedasticity functions for the whole temperature range, and present them in a 2D plot in Fig. 7. Such a scedasticity 2D pattern clearly reveals the reactivity pathway during the materials formation process. Traditionally, one can easily capture big changes of strong reflections in the diffraction patterns or large features in other spectra. Often subtle and important changes are submerged among the strong initial peaks, making them hardly distinguishable. The scedasticity will leach out unchanged parts and extrude active features during reactions. It is of particular importance when multiple phases coexisting, and some new or unknown intermediate phases forming during material synthesis and chemical reaction.

Conclusion

In summary, we have developed a generalized method to extract critical information from series of ranked correlated data, which is based on correlation functions and statistical scedasticity formalism. The method is applicable to all types of spectra evolving as a function of any arbitrary parameter. It is demonstrated in the in-situ synchrotron XRD study of the Li-ion battery cathode materials synthesis that the method can capture subtle spectrum change. Furthermore, the scedasticity function can track down which phases contribute the most in the change, reflecting a clear reactivity pathway. The methods described here can be readily used for on-the-

fly data analysis during various in-situ operando experiments in order to quickly evaluate and optimize experimental conditions, as well as for post data analysis and large data mining where considerable amount of data hinders the feasibility of the investigation through point-by-point inspection.

Acknowledgment

Research at Argonne National Laboratory was funded by U.S. Department of Energy, Vehicle Technologies Office. Argonne National Laboratory is operated for the U.S. Department of Energy by UChicago Argonne, LLC, under contract DE-AC02-06CH11357. The authors also acknowledge the use of the Advanced Photon Source of Argonne National Laboratory supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences.

References

- [1] Y.K. Sun, S. T. Myung, B.-C. Park, J. Prakash, I. Belharouak and K. Amine; Nat. Mater. 8 (2009) 320.
- [2] Y.K. Sun, Z. Chen, H.J. Noh, D.J. Lee, H.G. Jung, Y. Ren, S. Wang, C. S. Yoon, S.T. Myung and K. Amine; Nat. Mater. 11 (2012) 942.
- [3] Z. Chen, Y. Qin, K. Amine and Y.K. Sun; J. Mater. Chem. 20 (2010) 7606.
- [4] F. Zhou, W. Luo, X. Zhao and J. R. Dahn; J. Electrochem. Soc. 156 (2009) A 917.
- [5] H. Deng, I. Belharouak, Y.K. Sun and K. Amine; J. Mater. Chem. 19 (2009) 4510.
- [6] W. Luo, X. Li and J. R. Dahn; Chem. of Mater. 22 (2010) 5065.
- [7] W. Luo, F. Zhuo, X. Zhao, Z. Lu, X. Li and J. R. Dahn; Chem. of Mater. 22 (2010) 1164.
- [8] J. L. Shui, J. S. Okasinski, P. Kenesei, H. A. Dobbs, D. Z., J. D. Almer and D. J. Liu; Nature Comm. 4 (2013) 2255.
- [9] J. E. Katz, X. Zhang, K. Attenkofer, K. W. Chapman, C. Frandsen, P. Zarzycki, K. M. Rosso, R. W. Falcone, G. A. Waychunas, B. Gilbert; Science 337 (2012) 1200.
- [10] J. Wang, Y. C. Karen, C. Wiegart and J. Wang; Chem. Comm. 49 (2009) 6480.
- [11] J. Chen and J. Graetz; Am. Chem. Soc. Materials and Interfaces 3 (2011) 1380.
- [12] Z. Chen, Y. Ren, Y. Qin, H. Wu, S. Ma, J. Ren, X. He, Y. K. Sun and K. Amine; J. Mater. Chem. 21 (2011) 5604.
- [13] S. A. Wicker and E. H. W., Jr; Inorg. Chem. 52 (2013) 1772.
- [14] Q. Liu, H. He, Z. F. Li, Y. Liu, Y. Ren, W. Lu, J. Lu, E. A. Stach and J. Xie; Am. Chem. Soc. Applied Materials and Interfaces 6 (2014) 3282.
- [15] Z. Chen, Y. Ren, A. N. Jansen, C. K. Lin, W. Weng and K. Amine; Nat. Comm. 4 (2013) 1513.

- [16] Z. Chen, Y. Ren, E. Lee, C. Johnson, Y. Qin and K. Amine; Mater. Views 3 (2013) 729.
- [17] Z. Chen, Y. Qin, Y. Ren, W. Lu, C. Orendorff, E. P. Roth and K. Amine; Energy and Environmental Science 4 (2011) 4023.
- [18] Y. Cui, A. Abouimrane, J. Lu, T. Bolin, Y. Ren, W. Weng, C. Sun, V. A. Maroni, S. M. Heald and K. Amine; J. Am. Chem. Soc 135 (2013) 8047.
- [19] Y. Piao, Y. Qin, Y. Ren, S. M. H., C. Sun, D. Zhou, B. J. Polzin, S. E. Trask, K. Amine, Y. Wei, G. Chen, I. Bloom and Z. Chen; Phys. Chem. Chem. Phys. 16 (2014) 3254.
- [20] Y. Kan, Y. Hu, J. Croy, Y. Ren, C. J. Sun, S. M. Heald, J. Boreno, I. Bloom, Z. Chen; J. Power Sources 266 (2014) 341.
- [21] I. Noda; J. Am. Chem. Soc. 111 (1989) 8116.
- [22] I. Noda; App. Spec. 44 (1990) 550.
- [23] I. Noda; App. Spec. 47 (1993) 1329.
- [24] P. W. Faguy and N. Balachander, Analytical Comm. 34 (1997) 405.

Figure Captions:

Figure 1. Temperature dependent HEXRD patterns during the solid state synthesis of $\text{Li}(\text{Ni}_{0.7}\text{Mn}_{0.15}\text{Co}_{0.15})\text{O}_2$ using $(\text{Ni}_{0.7}\text{Mn}_{0.15}\text{Co}_{0.15})\text{CO}_3$ and Li_2CO_3 as starting materials.

Figure 2. Two dimensional covariance map of the HEXRD patterns.

Figure 3. Temperature dependent covariance matrix diagonals (i,i+1) in comparison to Pearson's correlation. The inset is a zoom-in of the 1st diagonal covariance and Pearson's correlation between 400°C and 800°C.

Figure 4. Results obtained from the Rietveld method: a) observed and calculated diffraction patterns at 450 C of the three phases; and temperature dependence of b) weight fraction of the three phases, c) lattice parameters of the layered and rock salt phases, d) Li concentration in cathode materials with the layered structure.

Figure 5. Diffraction patterns at 548 and 550 C.

Figure 6. A sordasticity function of the two diffraction patterns at 548 and 550C. Inset is a rescaled-colors zoom-in of the 2D diffraction.

Figure 7. 2D sordasticity pattern covering the whole temperature range.

Figures:

Figure 1, Aoun et al.

Figure 2, Aoun et al.

Figure 3, Aoun et al.

Figure 4, Aoun et al.

Figure 5, Aoun et al.

Figure 6, Aoun et al.

Figure 7, Aoun et al.