JOURNAL OF APPLIED CRYSTALLOGRAPHY

Volume 54 (2021)

Supporting information for article:

Validation of non-negative matrix factorization for rapid assessment of large sets of atomic pair distribution function data

Chia-Hao Liu, Christopher J. Wright, Ran Gu, Sasaank Bandi, Allison Wustrow, Paul K. Todd, Daniel O'Nolan, Michelle L. Beauvais, James R. Neilson, Peter J. Chupas, Karena W. Chapman and Simon J. L. Billinge

Validation of non-negative matrix factorization for rapid assessment of large sets of atomic pair-distribution function (PDF) data: Supplementary Information

Chia-Hao Liu,[a] Christopher J. Wright,[a] Ran Gu,[a] Sasaank Bandi,[a] Allison Wustrow,[b] Paul K. Todd,[b] Daniel O'Nolan,[c] Michelle L. Beauvais,[c] James R. Neilson,[b] Peter J. Chupas,[c] Karena W. Chapman[c] and Simon J. L. Billinge[a,d]

[a]*Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY 10016, USA,* [b]*Department of Chemistry, Colorado State University, Fort Collins, Colorado 80523, USA,* [c]*Department of Chemistry, Stony Brook University, Stony Brook, New York 11794, USA, and* [d]*Condensed Matter Physics and Materials Science Department, Brookhaven National Laboratory,Upton, New York 11973, USA*

Table S1: Parameters used to calculate PDFs from atomic structures for the ground-truth dataset. All parameters follow the same definitions as in (Farrow et al., 2007).

Parameter	Value
r_{min} (Å)	1.5
r_{max} (Å)	20.0
Q_{max} (Å^{-1})	25.0
Q_{min} (Å^{-1})	0.1
Q_{damp} (Å^{-1})	0.1
Q_{broad} (Å^{-1})	0.04

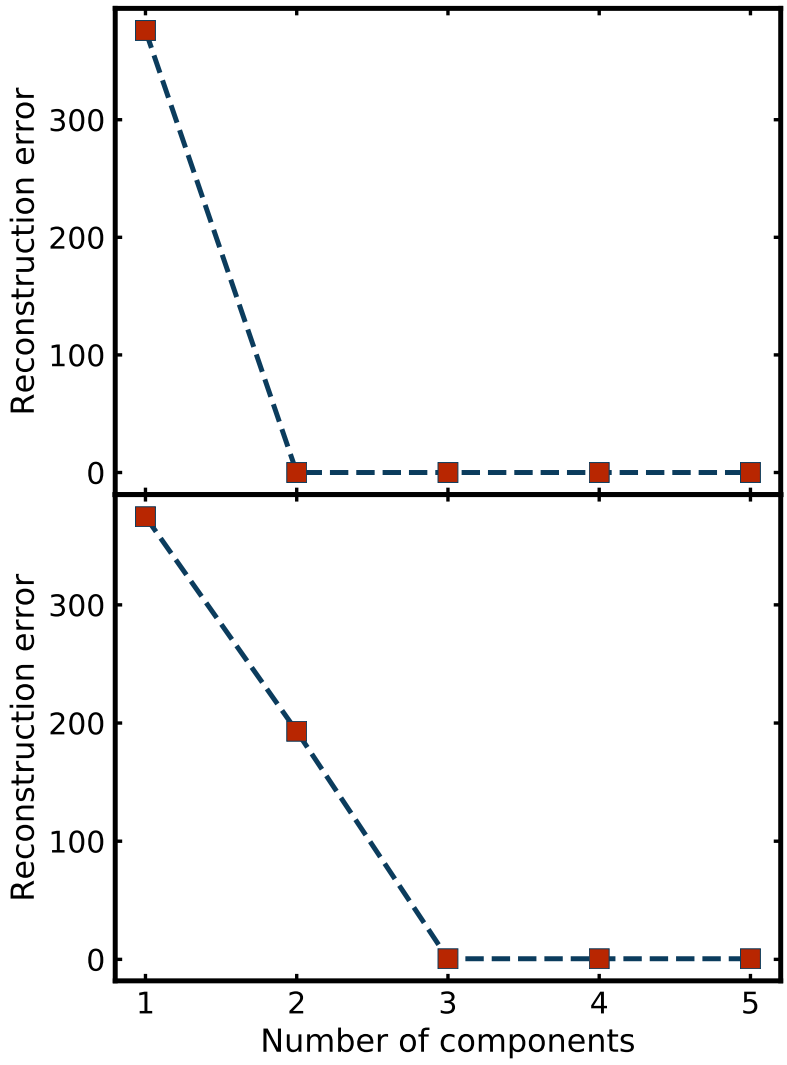

Fig. S1. The reconstruction error (defined in Eq. 4 in the main text) v.s. the number of components used for PCA (top) and NMF (bottom) algorithms on the simulated data.

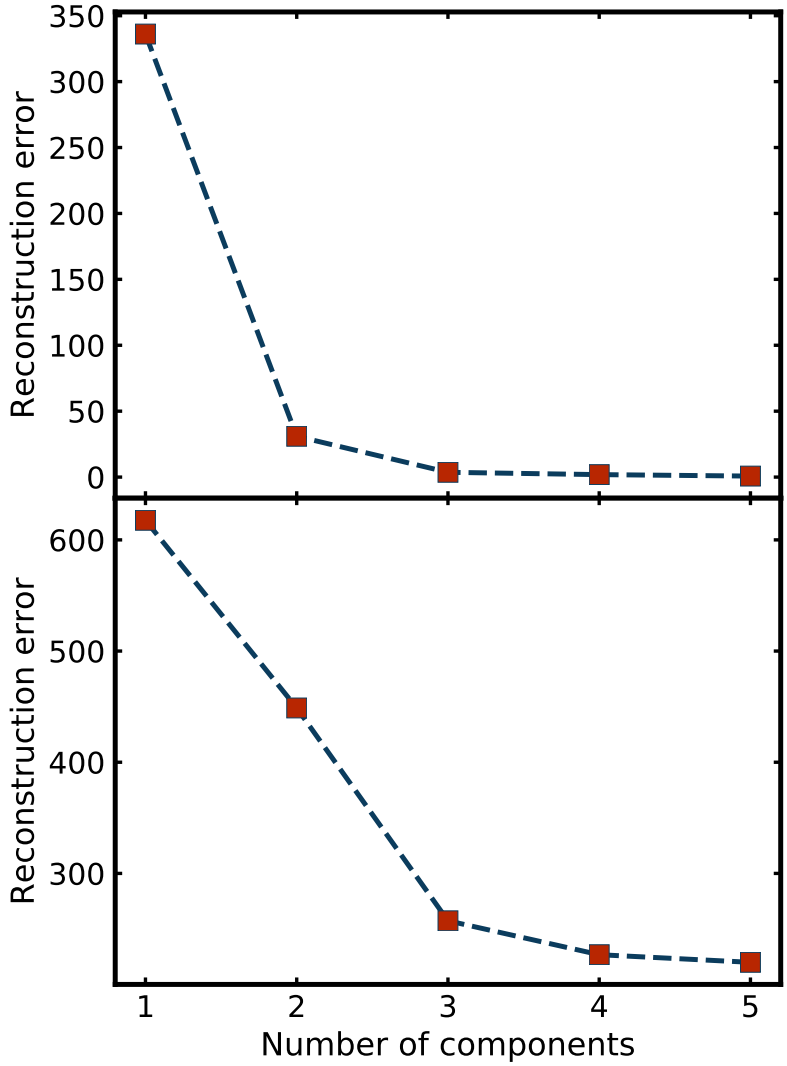

Fig. S2. The reconstruction error (defined in Eq. 4 in the main text) v.s. the number of components used for PCA (top) and NMF (bottom) algorithms on the experiment data.

References

Farrow, C. L., Juhás, P., Liu, J., Bryndin, D., Božin, E. S., Bloch, J., Proffen, T. & Billinge, S. J. L. (2007). J. Phys: Condens. Mat. **19**, 335219.