Nonnegative Matrix Factorization as a Tool for Improved Event Detection in Active Region Cores

SPD Student Poster Award Candidate

Will T. Barnes & Stephen J. Bradshaw

Department of Physics and Astronomy, Rice University

Introduction

- Determining the frequency of energy release in active regions (ARs) is a critical step toward learning how the solar corona is powered.
- Vast amounts of data $\rightarrow$ need for improved techniques for identifying impulsive heating signatures
- **Goal:** Predict heating frequency in AR cores using nonnegative matrix factorization (NMF) to analyze intensity fluctuations in forward-modeled pixel-averaged light curves

Nanoflare Signatures

- Nanoflares [5] are a strong candidate for the mechanism behind coronal heating
- Observational signatures difficult to detect$\rightarrow$ short timescale, non-equilibrium ionization, etc.
- Single peak may be convolution of many nanoflare signatures; counting methods not able to resolve these sources.

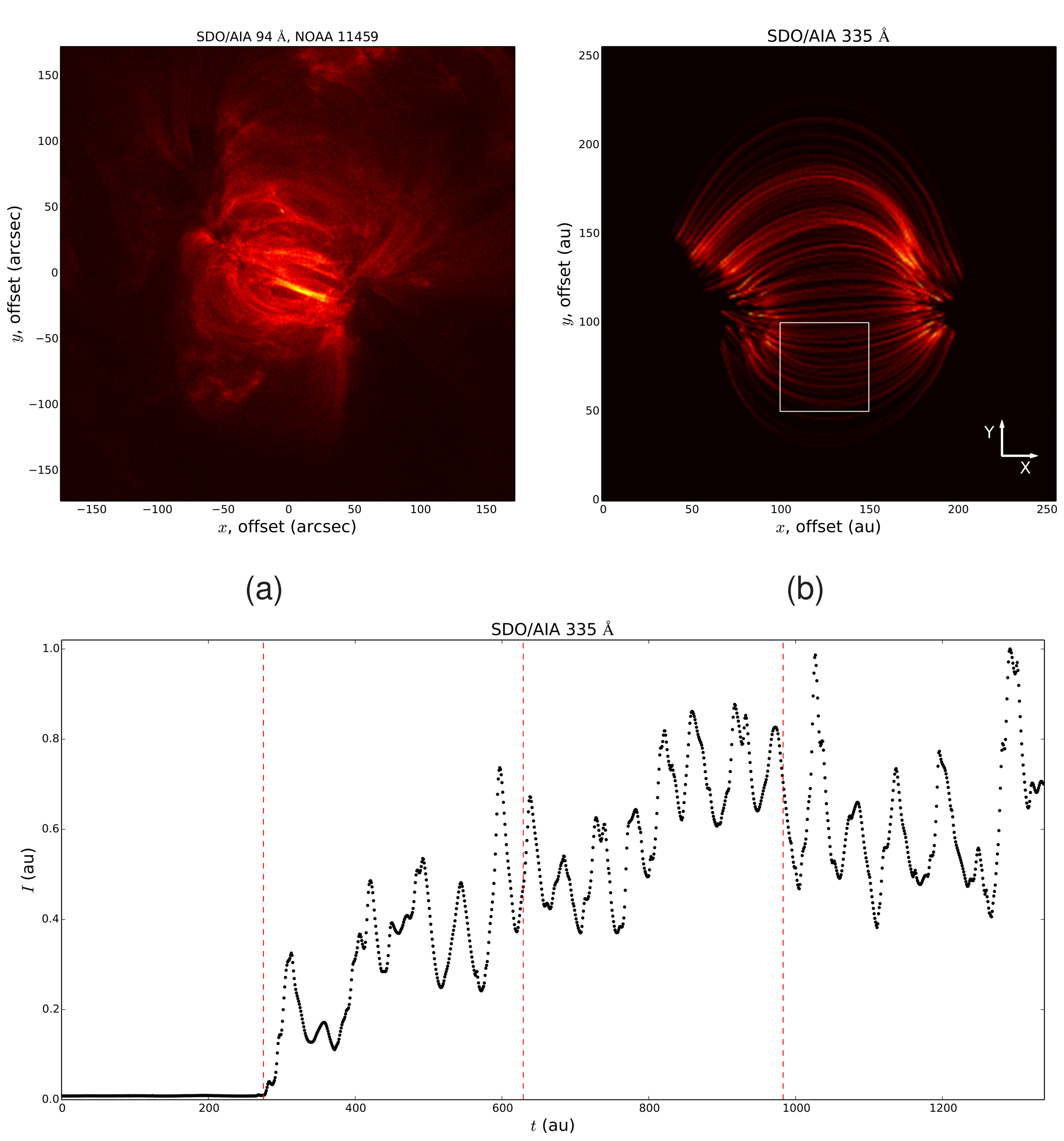

Figure 1: 1(a) SDO/AIA 94 Å image of AR NOAA 11459 from 22 May 2012, 00:02 UT. 1(b) Forward-modeled AR as seen by SDO/AIA 335 Å subject to a total of $\sim$4000 nanoflares. The white box shows the region over which the pixel-averaged time series was calculated. 1(c) Normalized time series calculated from pixel intensity inside white box of 1(b). Red lines show the different sections over which the factorization was performed.

Forward-modeling of AR Emission

- Forward-modeled emission provides ideal proof of concept for event detection algorithms
- Synthesized emission produced using n and T profiles from 1D hydrodynamic model HYDRAD [1] coupled with forward modeling code [2, 6].
- Resulting intentisity can be mapped to 3D field line extrapolations [e.g. 8] and the observed pixel intensity inferred from the integrated line-of-sight volumetric emission

Advantages of NMF

- Active algorithm development and testing by resarchers in a variety of fields [e.g. computational neuroscience, see 4, 7]
- Requires no information about source signal except guessed number of sources.
- Can be applied to 2D AR image data from SDO/AIA

Nonnegative Matrix Factorization

- Powerful tool used for matrix deconvolution using unsupervised learning algorithms
- Let some matrix $T \in \mathbb{R}_+^{m\times n}$ represent the observation. Goal of NMF is to factorize

$$T \approx UV,$$

where $U \in \mathbb{R}_+^{m\times k}$, $V \in \mathbb{R}_+^{k\times n}$, and k is the guessed number of sources.
- To represent T as $A = UV$, minimize divergence metric, $d(T|A)$, such that $d(T|A) = 0$ implies $T = A$.
- For $d(T|A)$, use modified version of the Frobenius norm derived by Chen & Cichocki [3] that controls for sparsity and smoothness.
- Chen & Cichocki [3] show that the following update rules for U and V can be derived such that $d(T|A)$ is always decreasing,

$$U_{ij} \leftarrow U_{ij}\frac{\sum_t V_{jt}T_{it}/A_{it}}{\sum_t V_{jt}},$$

$$V_{jt} \leftarrow V_{jt}\frac{(U^TT)_{jt}}{(U^TA + \lambda_1 VQ)_{jt} + \lambda_2/n(\sum_{i,i\neq j}^k V_{it} - V_{jt})}.$$

- U and V are updated iteratively until some desired convergence in $d(T|A)$ is met or a maximum number of iterations reached.
- The constituent sources can then be calculated as $A_i = U_i^{m\times 1}V_i^{1\times n}$ where $i = 1\dots k$ and $A = \sum_i A_i$.

Test Cases–1D & 2D

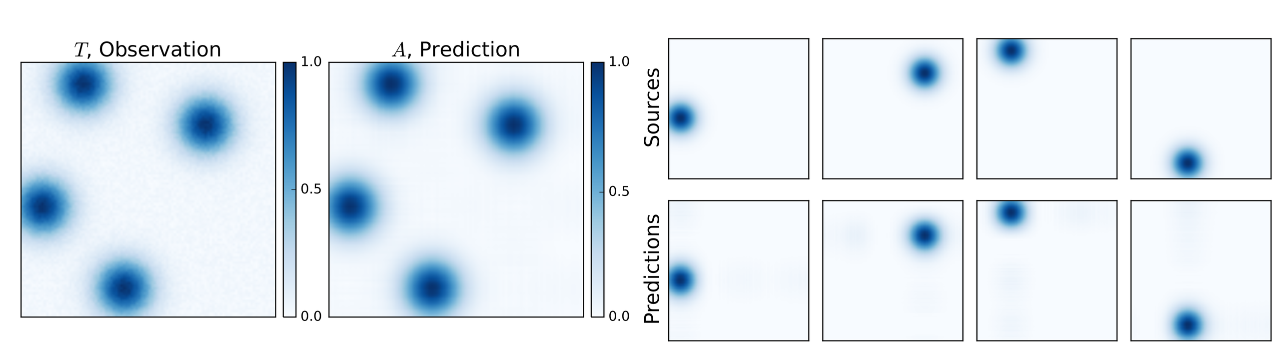

Figure 2: Example factorization for 4 gaussian sources with added noise. Sources are placed randomly on a 9×9 grid. All axes are in arbitrary units with amplitude scaled to unity.

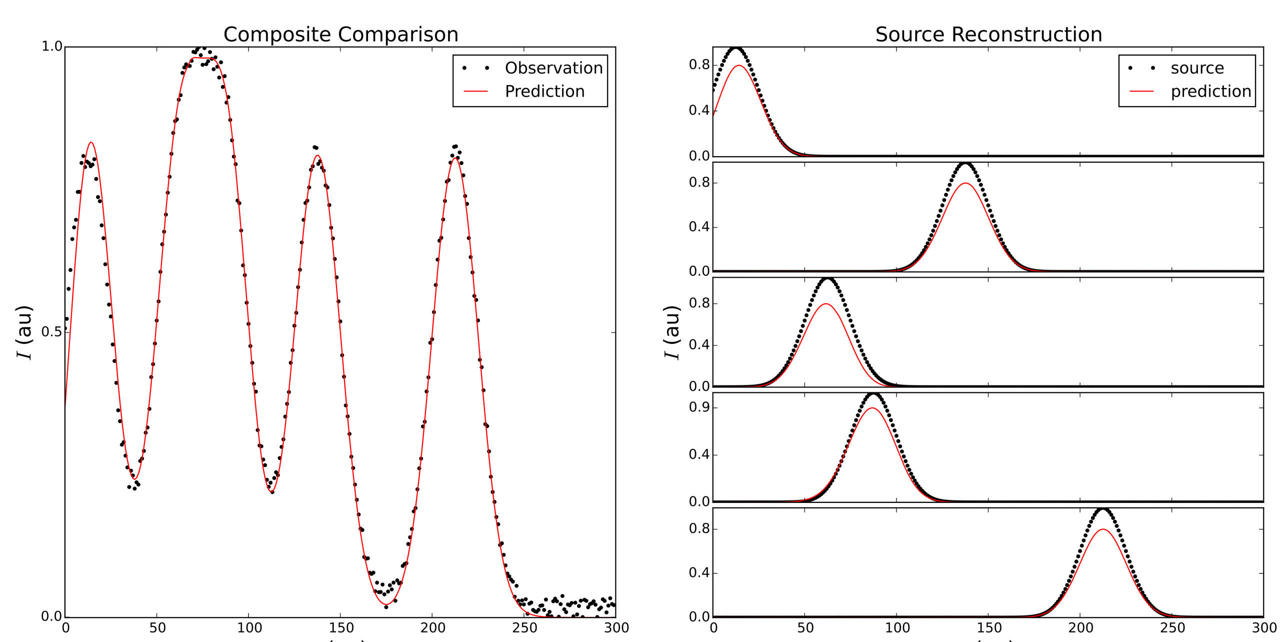

Figure 3: Example factorization for 5 gaussian sources with added noise. Note that NMF is able to recover the two heavily overlapping peaks near $t = 75$. Axes are in arbitrary units with amplitude scaled to unity.

Results

- NMF algorithm applied to forward-modeled SDO/AIA pixel-averaged time series (Fig. 1(c)) for $k \in [10, 50]$
- In order to estimate number of true sources, fit $d(k) = \alpha e^{-k/\tau_k}$ to $d(T|A)$ as a function of k such that the e-folding length τ_k is our estimate of the number of sources.

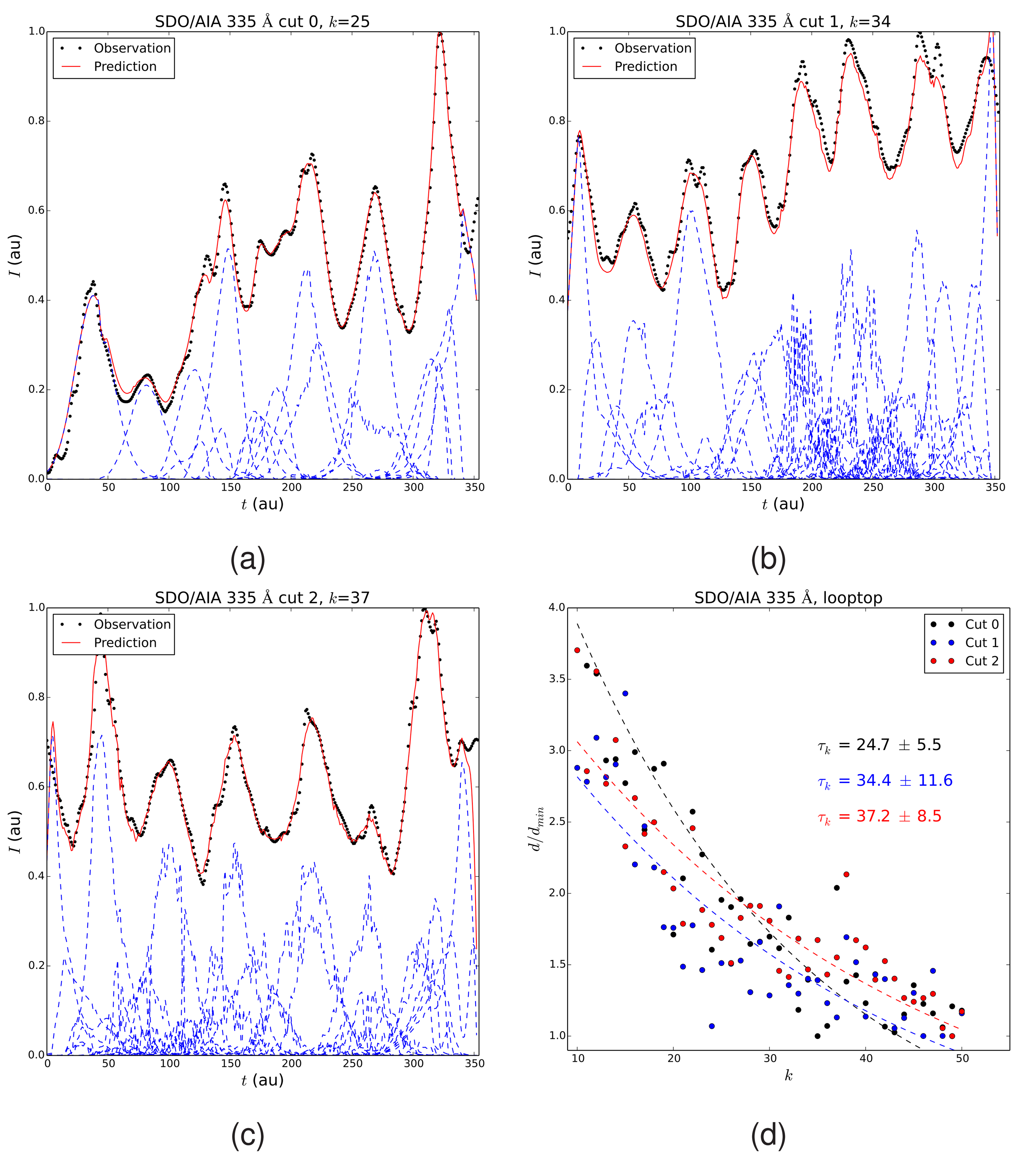

Figure 4: 4(a)-4(c) show the total reconstructed light curve plus the estimated underlying sources (in blue) for the light curve sections shown in Fig. 1(c). 4(d) shows the final value of $d(T|A)$ (normalized to $\min(d(T|A))$ as a function of k; dotted lines show the exponential fit.

Conclusions & Future Work

- Expected number of events per cut $= A_{\text{white box}}/A_{AR} \times N_{\text{total}}(= 3448)/3 \approx 86 > N_{\text{predicted}}$ for all sections of the 335 Å light curve.
- As k increases, $d(T|A)$ tends to decrease regardless of number of true sources; testing with gaussian sources(i.e. Fig. 3) will help to address this effect
- Single-pixel and filtered forward-modeled emission [see 9] will provide more conclusive tests.
- True advantage of NMF will be seen in application to 2D AR images (i.e. Fig. 1); interesting and exciting alternative to traditional detection methods.

References

[1] Bradshaw, S. J., & Cargill, P. J. 2013, The Astrophysical Journal, 770, 12
[2] Bradshaw, S. J., & Klimchuk, J. A. 2011, The Astrophysical Journal Supplement Series, 194, 26
[3] Chen, Z., & Cichocki, A. 2005, in Laboratory for Advanced Brain Signal Processing, RIKEN, Tech. Rep
[4] Cichocki, A., Zdunek, R., Phan, A. H., & Amari, S.-i. 2009, Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation (Wiley Publishing)
[5] Parker, E. N. 1988, The Astrophysical Journal, 330, 474
[6] Reep, J. W., Bradshaw, S. J., & Klimchuk, J. A. 2013, The Astrophysical Journal, 764, 193
[7] Soelter, J., Schumacher, J., Spors, H., & Schmuker, M. 2014, NeuroImage, 98, 279
[8] Warren, H. P., & Winebarger, A. R. 2007, The Astrophysical Journal, 666, 1245
[9] Warren, H. P., Winebarger, A. R., & Brooks, D. H. 2012, The Astrophysical Journal, 759, 141

Department of Physics and Astronomy - Rice University - Houston, TX

wtbarnes Mail: will.t.barnes@rice.edu