Non-Negative Matrix Factorisation (NNMF) for Recovery of Time Traces from Overlapping Signals

Siddharth Doshi, and Chance Ornelas-Skarin

Abstract

NNMF has been widely used for recovery of noisy time varying signals, such as fluorescent traces from biological tissue. Recently the technique has also been applied in extracting time traces of signals that have been scattered and mixed by thick tissue, without the need to recover their original spatial footprints. We explore this further through the use of NNMF for recovery of time traces from noisy, overlapping simulated video data. We show NNMF can recover time traces from overlapping signals up till certain thresholds of noise and overlap degree, but draw attention to circumstances, including complex modulation patterns and high overlap degrees, where performance is significantly poorer.

Index Terms

Computational Photography

✦

1 Introduction

Recovery of signals from through highly scattering tissue (e.g. time varying signals such as fluorescent traces from biological tissue) is an important problem. A particular restricted case is when we are interested in simply extracting time traces of scattered signals without explicitly recovering their original (pre-scattered) footprint. This can be thought of as a blind source problem. A powerful approach is NNMF, used for separation of spatial and temporal footprints to allow for denoising by separating signal traces from background [2]. Recently this approach was used to recover the temporal variations of signal fluorescent markers under scattering tissue [1].

NNMF is a form of unsupervised learning similar to PCA where the components of some dataset are extracted and ranked based on their variance [3]. The highest ranked components describe the most characteristic components of the data based on the way the components were extracted. NNMF specifically extracts these components with the constraint that the components must be non-negative. Since many measurements are composed of purely positive values, this constraint makes sense to have for those measurements.

We illustrate in Fig 1. how a vectorized video dataset can be factorized into a matrix of weights representing a spatial component, e.g. for each "cell" and a temporally varying component, for the associated time trace of each cell.

2 Related Work

NNMF, and modified versions, have been widely used for a range of signal processing tasks in many application domains involving some decomposition of data which are expected to have non-negative components that add together. In a landmark paper, NNMF is compared to principal component analysis and vector quantization as a way creating a decomposition basis for representing the parts of a human face [3]. They showed that the non-negativity constraint, that all the values in its basis are positive, leads to more powerful, intuitive and interpretable decompositions of parts of faces than techniques such as PCA. Composing a face from a basis as a purely additive process gave better outcomes than when components were free to be negative. They showed that the technique works well for problems where a parts-based representation is desired. In the specific domain of imaging, NNMF has found wide use in the de-mixing and de-noising of time varying biological imaging data, such as calcium imaging data from neuronal signals. One widely used version is constrained non-negative matrix factorisation (CNMF), used recovering temporal and spatial signals from data captured from overlapping neurons [2]. They propose two nonlinear matrix factorization strategies, NNMF and multilevel sparse matrix factorization. They start by using a constrained deconvolution approach where the sparsest neural activity signal is desired. They embed this constrained deconvolution in a constrained NNMF framework that considers the properties of the expected signals as constraints. They can use this to successfully capture the spatial signature of neural signals and their temporal signature with little human intervention. Most recently, this technique has been shown to work in the even more extreme use case of recovering time varying traces through thick scattering tissue, where the light is completely mixed

- *S.Doshi and C.Ornelas-Skarin are with the School of Engineering, Stanford University, Stanford, CA, 30332.*
-

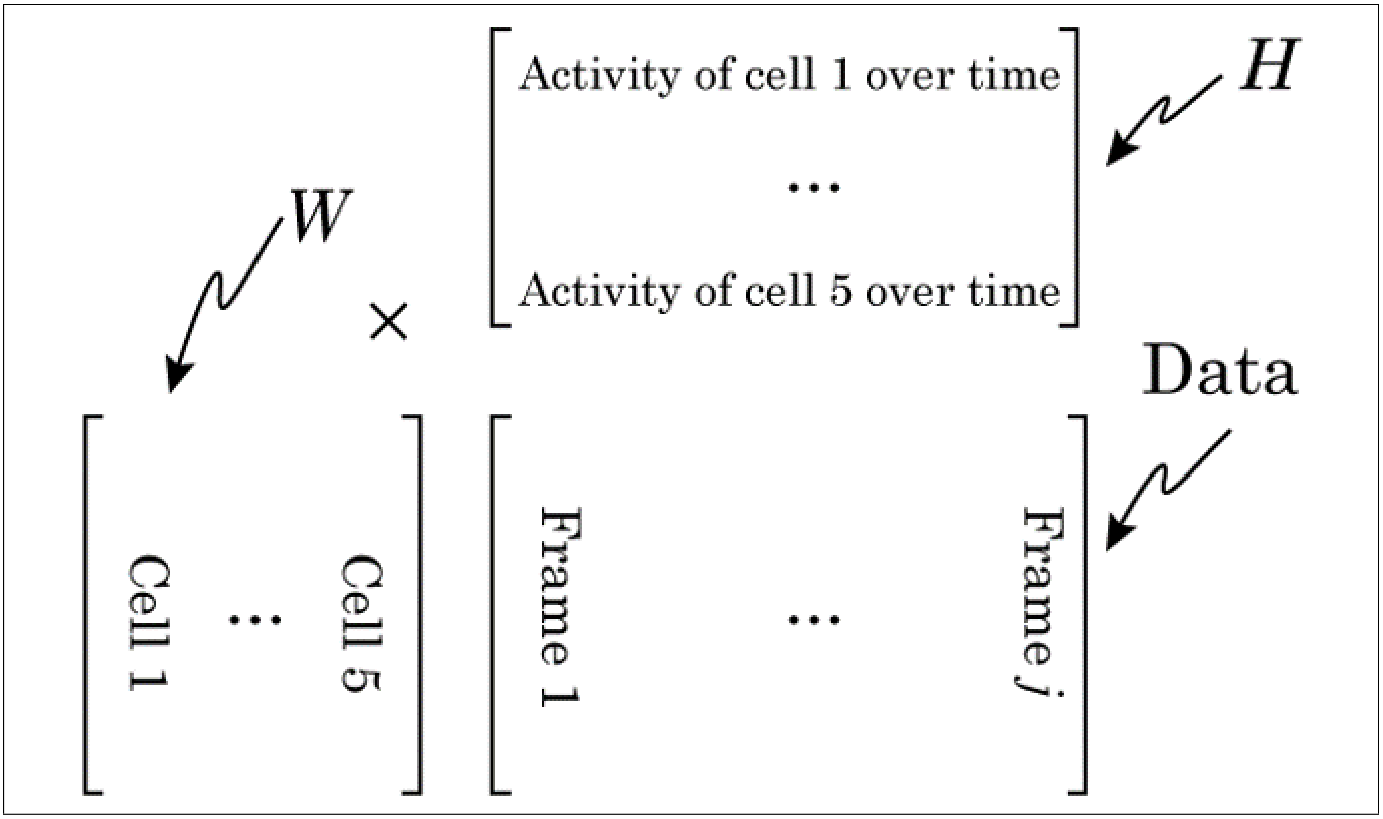

Fig. 1. Separation of dynamic trace data into spatial and temporal components, figure adapted from [4].

and scattered and only speckle patterns associated with the scattering are imaged. This could have significant application in areas such as neuronal imaging, where in many cases the time traces of the images are the only important pieces of information, and spatial information (in this case the imaged speckle patterns associated with each fluorescent source) can be discarded. In this first work [1] they use an experimental set-up where fluorescent beads are used to replicate the signatures of signals that might be detected from neuron activity. They do this by activating the beads with a spatial light modulator and allowing it to scatter through a mouse skull and record the resulting speckle patterns on a CMOS camera. Then they use NNMF to find a solution for the temporal and spatial fingerprints of the beads. They note that the randomness in the scattering from the tissue helps regularize the problem. They were able to successfully separate the signals of buried fluorescent beads after the scattered through the mouse skull.

3 Proposed Method

We expect that for imaging through highly scattering tissue to be feasible, distinct speckle patterns are needed for each source (e.g. bead) to be imaged. This introduces requirements such as the use of coherent light to create spatially varying complex speckles, and that the beads used are separated further than the memory effect of the tissue. Incoherent light, which would create diffuse, blurred and uniform overlapping patterns of light that aren't distinct could be much more difficult to extract information from. To begin to investigate this point, we will explore the separation of signals from a dataset simulating scattering of incoherent light from distinct time varying signals through thick tissue, as a function of noise and degree of overlap.

3.1 Creation of Simulated Data

Create simulated "video" dataset by modelling the pattern of light from a single source region after scattering as a Gaussian spots, and then creating an image frame by superimposing many Gaussians with varying degrees of overlap. We then apply different time varying intensity traces to each of these Gaussians and apply noise. We implement the NNMF algorithm and benchmark its accuracy in recovering our original "ground truth" time intensity traces under various noise conditions.

3.1.1 Investigation Plan

We generate example datasets with various degrees of spatial overlap (Fig 2).

We then apply different time modulation signals to different Gaussians. We show two types of modulation, one with sinusoidal modulation at different frequencies (Fig 3a.), and different "neural-like" spike data (Fig 3b.).

We also vary the noise level and apply different types of noise (uniform and normal) in a static and time varying fashion. Illustration of the noising is shown (Fig 4).

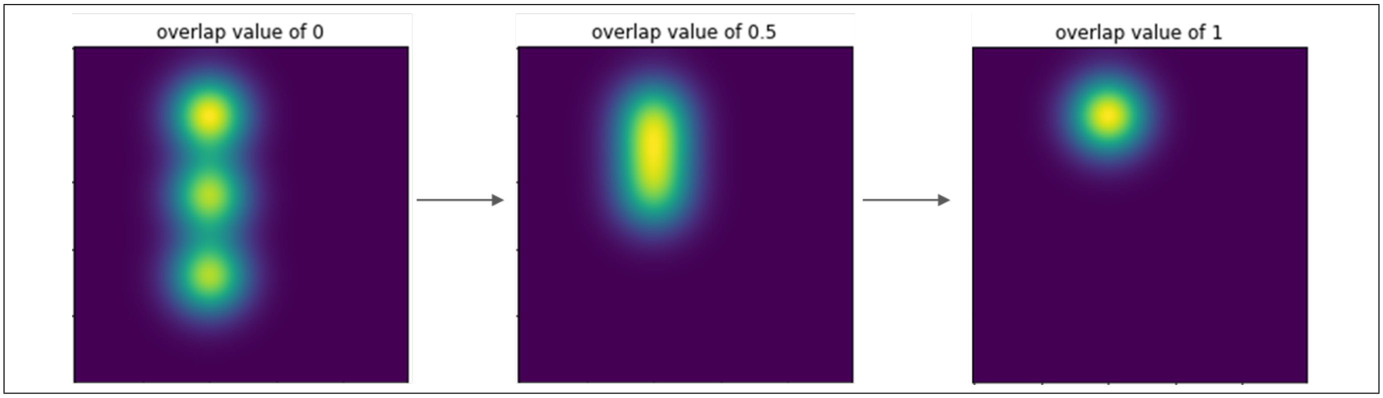

Fig. 2. Degrees of overlap in simulated data.

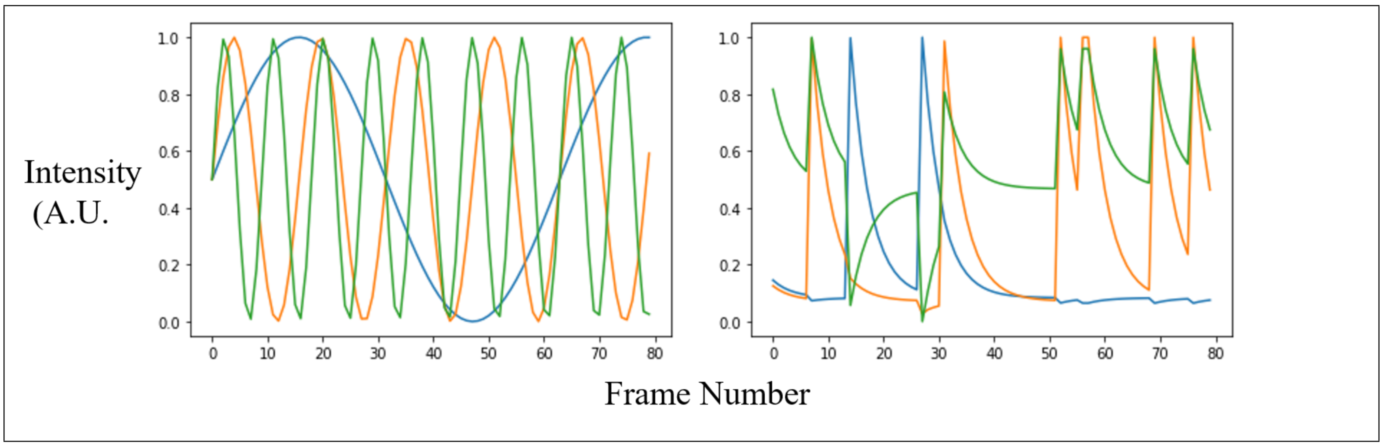

Fig. 3. a) (left). Examples of sinusoidal time traces, b) (right). Neural ”spike train” time traces.

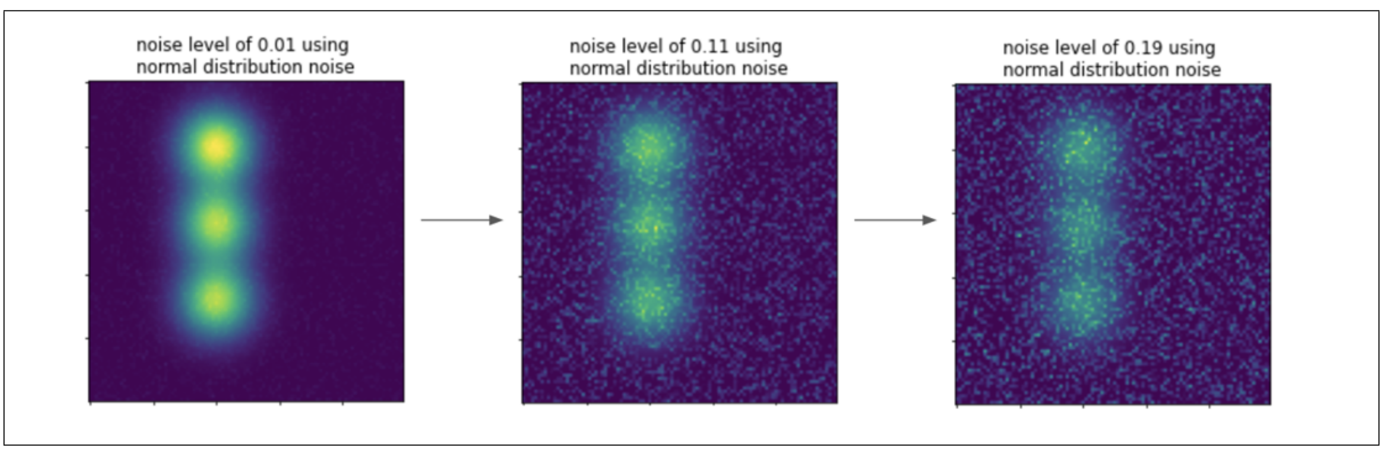

Fig. 4. Noise applied to simulated data.

4 Experimental Results

4.0.1 Recovery of randomly positioned sinusoidally modulated traces

We first show that in Fig 5. that NNMF allows for recovery of ground truth traces with high fidelity from sinusoidal applied patterns, with Gaussian spots applied randomly and without overlap.

4.0.2 Recovery of modulated traces at various overlap degrees as a function of noise

We then investigate how the reconstruction fidelity varies as a function of noise for different noise types, with various degrees of overlap, again for sinusoidal signals. We quantify the reconstruction by taking the L2 norm of the difference between recovered traces and the ground truth.

Results are show for sinusoidal modulation with uniform noise in Fig 6 and with normal noise in Fig 7.

We also show the results as a function of normal noise for ”spike” modulated data in Fig 8.

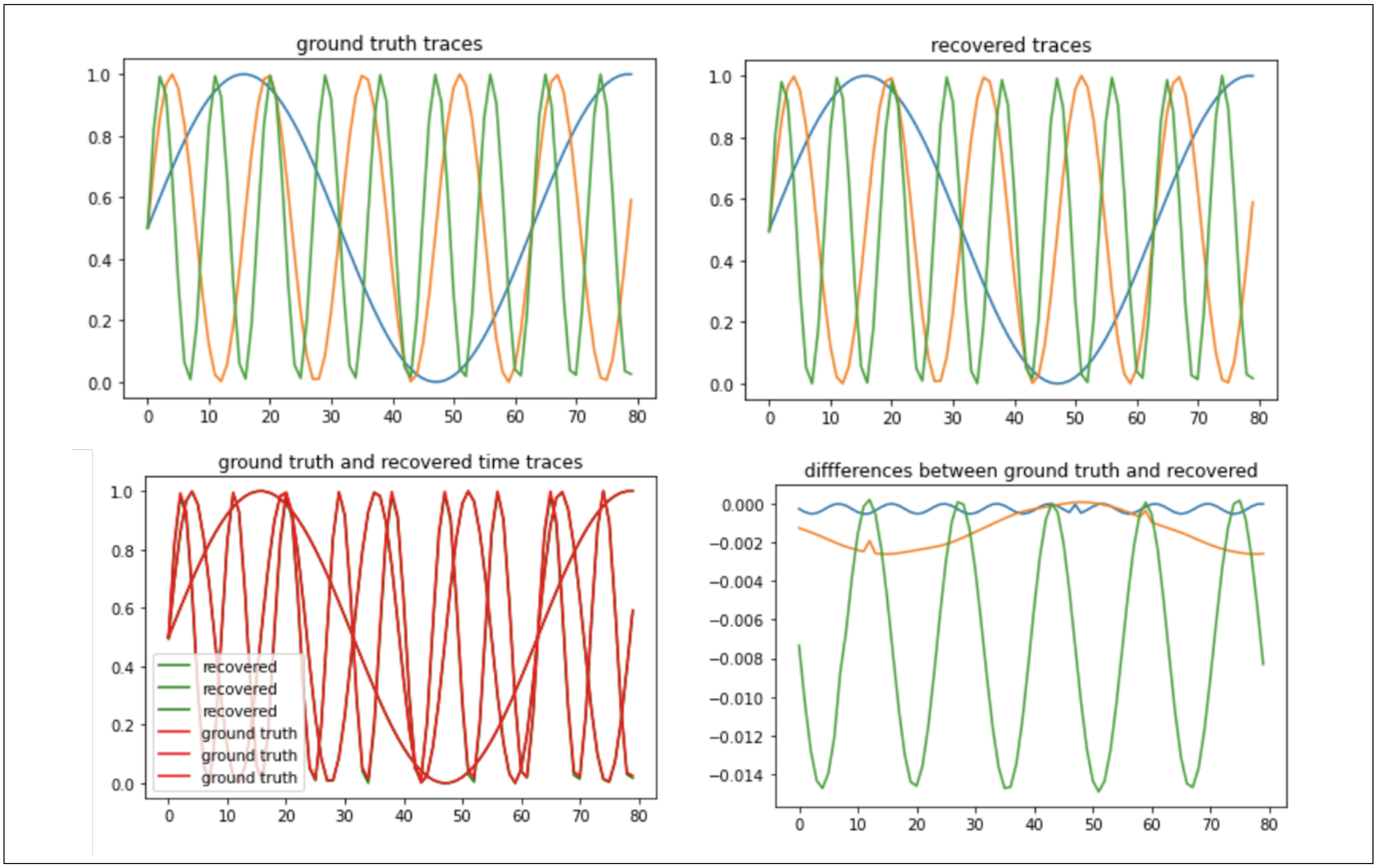

Fig. 5. Recovery of sinusoidally modulated traces.

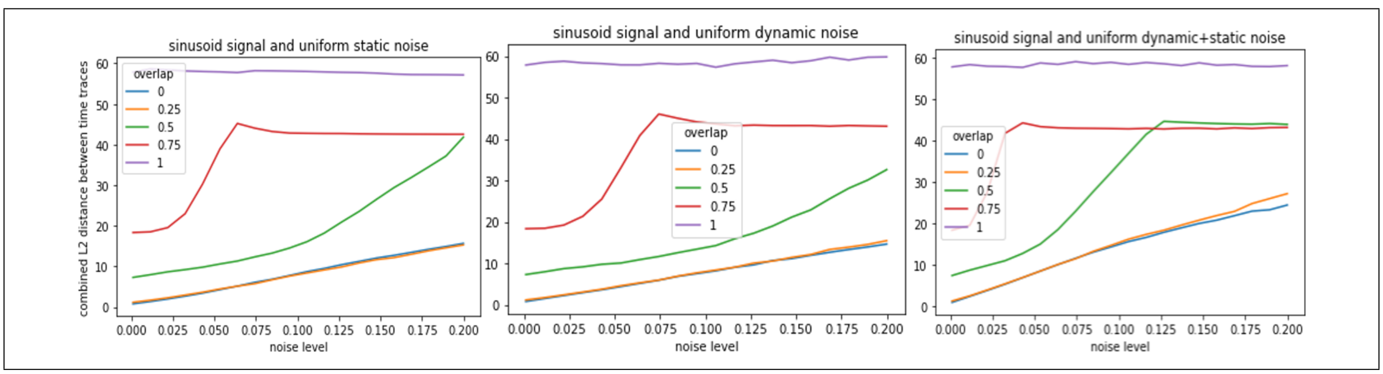

Fig. 6. Recovered fidelity (defined by L2 norm of difference between recovered and ground truth signal) as a function of uniform noise level for different overlap degrees.

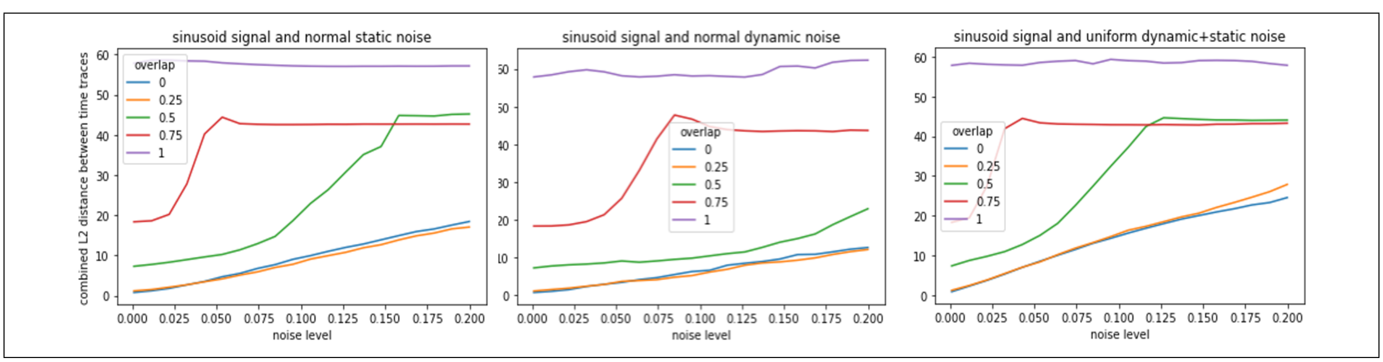

Fig. 7. Recovered fidelity (defined by L2 norm of difference between recovered and ground truth signal) as a function of normal noise level for different overlap degrees.

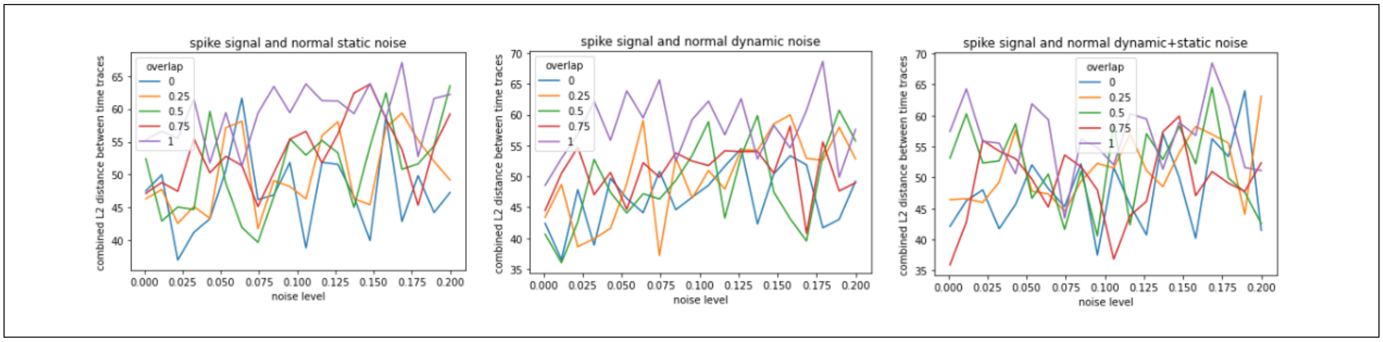

Fig. 8. Recovered spike data fidelity (defined by L2 norm of difference between recovered and ground truth signal) as a function of normal noise level for different overlap degrees.

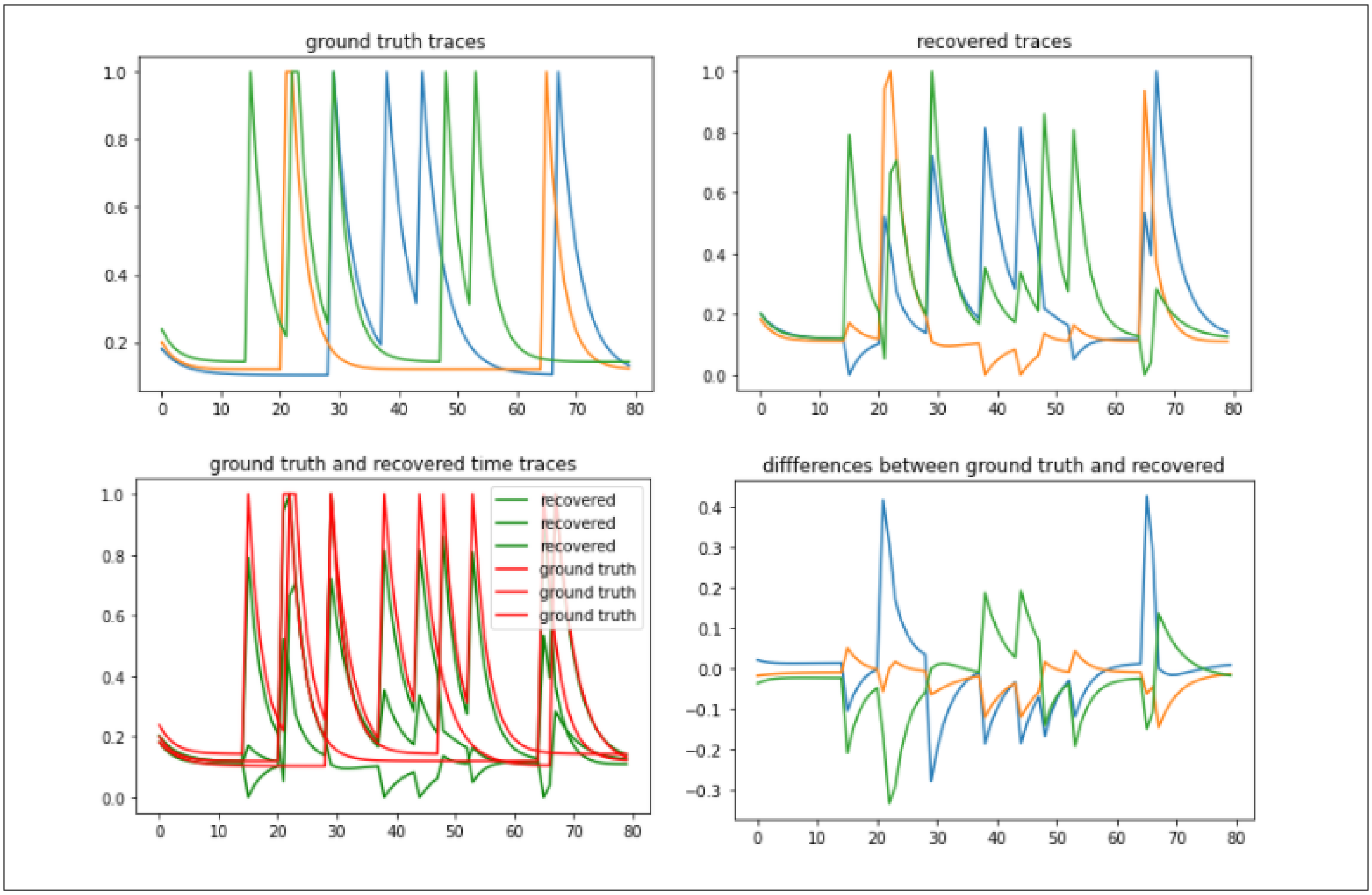

Fig. 9. Example of recovery of Spike data Traces

5 Discussion

We see that NMF allows for recovery of ground truth traces with high fidelity from simple sinusoidally applied patterns without overlap. When applied to sinusoidal time trace patterns, at low overlap, the signals can be recovered even in the presence of increasing noise, for all cases of static, dynamic and static and dynamic noise. Even at full overlap, the signals can be recovered at low noise conditions and the quality of the reconstructions decrease as noise increase, before hitting a plateau where they are too noisy to reconstruct. However, at full overlap, the reconstructions are poor even at low noise, indicating that the system cannot recover fully overlapping signals. However, for simple sinusoidal signals, it is highly promising that even “75 per cent” overlapped signals can be recovered in some noise. This implies that signals that aren’t fully overlapped (e.g. speckle patterns with partial overlap) can be reconstructed well, but if a scattering media causes full overlap of signals (e.g. of incoherent light diffusely scattering through a media), then reconstruction is not feasible. An interesting point for future investigation is whether reconstruction is possible for data where one gaussian is a enclosed within another gaussian (e.g. a small spot with time trace x inside a large spot with time trace y).

We note that on the spike data, recovery from an L2 perspective is overall quite poor for all types of noise and overlap. This is likely because our “L2 norm” fidelity criterion is not accurate at representing how well the system gets the spike timing right. Observing the actual ground truth and reconstructed traces visually, we see that the system can recover some spikes but misses some spikes, and gets amplitudes wrong. This indicates that recovery mostly occurs, but with some mistakes that show up as large errors in the L2 norm, indicating we need a better criterion to systematically study the effect of noise on the recovery. However, we note in general that the imperfect reconstructions (including the missed spikes and inclusion of spuerious spikes) indicates this method is not perfect for these complex signals, and perhaps that more complex frequency content signals are harder to reconstruct.

6 Conclusion

NNMF is a valuable technique for recovery of time traces from noisy data. It can recover time traces from overlapping signals up till certain thresholds of noise and overlap degree. It's performance on complex trace data can be further investigated, along with the possible use of alternative algorithms such as constrained non negative matrix factorization to improve fidelity (CNMF) [2].

7 References

[1] Moretti, C., Gigan, S. Readout of fluorescence functional signals through highly scattering tissue. Nat. Photonics 14, 361–364 (2020). https://doi.org/10.1038/s41566-020-0612-2

[2] Pnevmatikakis EA, Soudry D, Gao Y, Machado TA, Merel J, Pfau D, Reardon T, Mu Y, Lacefield C, Yang W, Ahrens M, Bruno R, Jessell TM, Peterka DS, Yuste R, Paninski L. Simultaneous Denoising, Deconvolution, and Demixing of Calcium Imaging Data. Neuron. 2016 Jan 20;89(2):285-99. doi: 10.1016/j.neuron.2015.11.037. Epub 2016 Jan 7. PMID: 26774160; PMCID: PMC4881387.https://www.cell.com/neuron/pdfExtended/S0896-6273(15)01084-3

[3] Lee, D., Seung, H. Learning the parts of objects by non-negative matrix factorization. Nature 401, 788–791 (1999). https://doi.org/10.1038/44565 [4] Connelly, B., 2022. Nonnegative Matrix Factorization for Dummies. — Bill Connelly. [online] Billconnelly.net. Available at: ¡http://www.billconnelly.net/?p=534¿ [Accessed 11 March 2022].