# MULTI-CHANNEL BLIND SOURCE SEPARATION OF SEISMIC ARRAY DATA

**Sheldon Deeny**

**Program in Applied Mathematics**

**University of Arizona**

**Tucson, AZ**

**Dr. Daniel Champion**

**Senior Principal Scientist**

**Nevada National Security Site**

**Las Vegas, NV**

**This work was done by Mission Support and Test Services, LLC, under Contract No. DE-NA0003624 with the U.S. Department of Energy. DOE/NV/03624--1255**

# TABLE OF CONTENTS

# ABSTRACT

Blind source separation (BSS) is the attempt to unmix a signal composed of multiple sources. It is blind because there is usually little to no prior knowledge about the total number of sources in the observed signal, how they are mixed, or anything else about the system in which the signals are recorded. The signals themselves may represent a time series, images, or higher dimensional data without any time component.

This research summary describes how different methods and algorithms were applied to a blind signal separation problem with multi-channel seismic array data. There are several considerations which are unique to this application, such as time offsets of signals between sensors and the presence of many nuisance signals. Two of the most common methods for source separation come from applying the principal component analysis (PCA) and nonnegative matrix factorization (NMF) of a data matrix $X$. After the introduction and background, we look at how these two methods were able to assist in source separation of the seismic data.

# INTRODUCTION AND PROBLEM BACKGROUND

The Advanced Data Analytics for Proliferation Detection (ADAPD) project is a joint venture between the Nevada National Security Site (NNSS) and five other national labs: Los Alamos, Pacific Northwest, Sandia, Lawrence Livermore, and Oak Ridge. It is focused on three "hard problems" related to detecting proliferation, which is the activity of nuclear capability development. One of the problems is being primarily researched at the NNSS.

The U1a underground facility (Fig. 1), located at the NNSS site north of Las Vegas, has been a model testing location for the collection of data and development of analytic capabilities. U1a attributes and activities include: government controlled restricted site, underground facility, large-scale science and engineering deployments, and mining activity, which are all of high importance in the development of proliferation detection capabilities. High quality analytics cannot be developed and tested without collecting high-quality data. To this end, the site is equipped with multi-modal sensors which are able to continually monitor and record various metrics: power usage, cyber & IT activity, and seismic-acoustic energy sources recorded from a surface array of 3-axis geophones. The latter is the data mode of interest in this research summary.

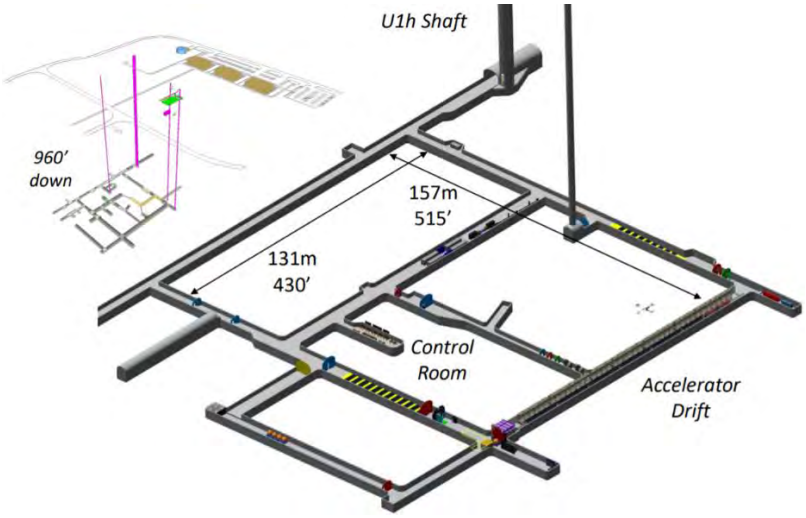


Figure 1: U1a Facility schematic

Several analytic projects using the seismic data have been proposed. They include: (1) sub-terranean time-difference-of-arrival (TDOA) multi-lateration; (2) mining related signature development; (3) underground facility support infrastructure signature development; (4) elevator shaft signature and predictive analytic; and (5) surface vehicle activity, timing, directionality, and entrance/exit signature development. The seismic geophone sensors are able to record all manner of sources which emit energy in the form of seismic and acoustic waves. The variety of human and machine activity at U1a, at the surface and underground, creates superimposed signals from different sources and different locations. For example, at the same time, a heavy door being closed, percussions from a rock bolting machine, ventilation shaft fan motors, and a forklift driving on a surface road all contribute to the recorded signal. Additionally, the complex is 293m below the surface and constructed entirely within unsaturated alluvium, composed mainly of tuffaceous sand, with alternating and interbedded layers of sandy gravel and cobbly sand. The challenge of robust signal detection through such a thick, attenuative medium using surface sensors cannot be understated.

In all of the aforementioned applications, it is a central task to be able to understand which types of signatures are coming from which possible sources. The data collected is noisy and contains a mixture of overlapping signatures, interference, attenuation of signals, nuisance tones, and time-domain offsets among the multiple sensors. As seen in the Fig. 2 spectrogram image, a variety of broad and narrow spectrum signals appear in a short time span. In order to develop specific signature detection and predictive analytics, we need to apply digital signal processing techniques to separate the mixture. In general, this presents itself as a *blind source separation* problem. In the context of machine learning, we can think of this as an unsupervised learning problem, in that we know very little or nothing about the signals themselves, i.e. labeled data is hard to come by.

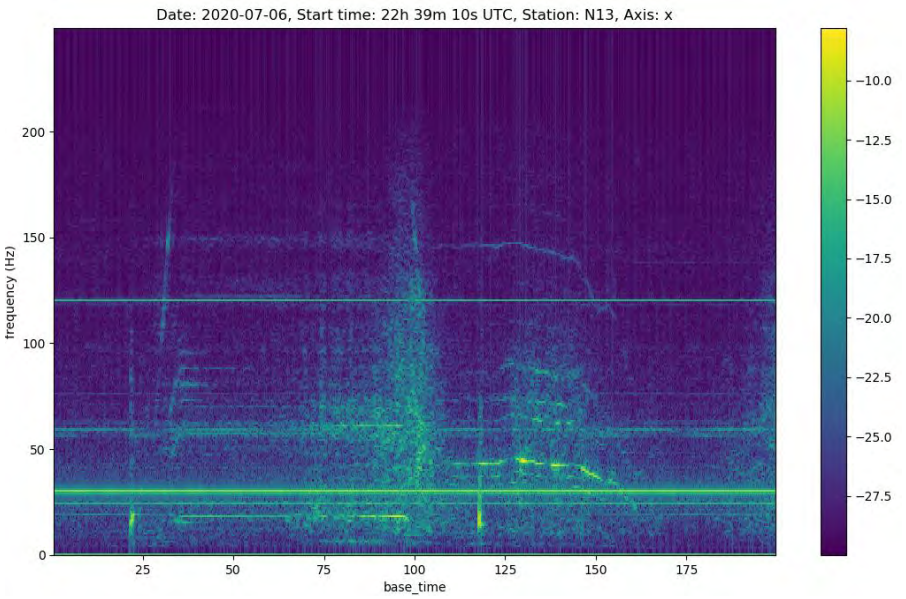


Figure 2: Spectrogram

In the literature, the most common example mentioned when introducing BSS is the *cocktail party problem*. This involves the task of "hearing" or "listening" to a conversation in a crowded cocktail party, where there are many simultaneous conversations, music, and other source noises being received by the "listener," who is tasked with focusing on a single conversation. The human listener, with two ears (sensors) is able to process this task by filtering out the distracting noises. The human brain is able to achieve source separation. For a computer, processing a digitally sampled, mixed signal is difficult. The cocktail party problem itself can be classified under the broader setting of auditory scene analysis. Digital signals can represent more than audio signals, of course. Separation may need to be applied seismic data (like in our case), images, EEG data, text documents, or any manner of time series data, to name a few examples.

In addition to source separation, we wish to achieve *multilateration*, that is, be able to locate a specific source once it has been separated from the mixed signal. This is a well-developed area used in guidance and surveillance systems. It is based on finding the intersection of hyperbola from timing offsets in synchronized sensors, where least squares methods can be used when there are more than 4 sensors/receivers used to locate a source in three dimensions. In our use-case, the sensor spacing, wave propagation speed, and wave media all contribute to offsets between sensors of the same source signal.

This report gives an overview of two methods which have been developed for the blind separation problem at U1a. BSS is highly underdetermined in most cases. It can be *monaural*, that is, used to separate mixed sources from a single sensor, or applied with many sensors. The work discussed herein was conducted during the student summer intern program at MSTS LLC., civilian contractor for the NNSS, under the guidance of Senior Scientist Dr. Daniel Champion. A portion of it had already been developed by Dr. Champion, and it is outlined first. This is followed by initial results and a discussion about how well the methods performed when used on real data.

# METHODS

**Mathematical Formulation**

In this section I will lay out the basic mathematical formulation and assumptions for source separation of the seismic array. The unknown sources are emitted from unknown locations and are mixed together as they propagate through the ground and air, subject to interference, reflection, echoes, and other physical factors.

We assume a linear system in the sense that at some time $t = 1, \dots, T$, the signal picked up by a sensor $X_i(t)$ for $i = 1,2, \dots, M$ is a linear combination of sources $S_j(t)$ for $j = 1,2, \dots, N$. For now, we ignore the typical 0 mean normally distributed error term $\epsilon_i$ added to the signal.

$$X_i(t) = a_{i1}S_1(t) + a_{i2}S_2(t) + \cdots + a_{iN}S_N(t) \qquad (1)$$

The $a_{ij}$'s are attenuation (or possibly amplification) factors which are determined by the nature of the mixing system. For $M$ sensors and $N$ sources, the data can be represented as a matrix $X$ which is the result of the mixing system matrix $A$ acting on the $N$ sources:

$$X = AS \qquad (2)$$

where $X \in \mathbb{R}^{M \times T}$, $A \in \mathbb{R}^{M \times N}$, and $S \in \mathbb{R}^{N \times T}$. That is, each row of $X$ is a digitally sampled signal from an individual sensor for $t = 1,2, \dots, T$, the entries in the matrix $A$ are the attenuation factors, and each row of $S$ is an individual source. Matrix $A$ can also be thought of as the *gain* matrix, which determines the weights of each source which make up a signal. Equation (2) can be written out fully as

$$\begin{bmatrix} x_1(1) & \cdots & x_1(T) \\ \vdots & \ddots & \vdots \\ x_M(1) & \cdots & x_M(T) \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{M1} & \cdots & a_{MN} \end{bmatrix} \begin{bmatrix} s_1(1) & \cdots & s_1(T) \\ \vdots & \ddots & \vdots \\ s_N(1) & \cdots & s_N(T) \end{bmatrix} \qquad (3)$$

It should also be mentioned that a simplifying assumption is the matrix $A$ is unchanging in time, that is, the mixing system is the same during the entire signal collection.

If matrix $A$ can be inverted, we can recover the sources: $S = A^{-1}X$. However, $A$ is in general unknown and unobservable, which makes the problem interesting. Blind source separation techniques attempt to estimate $A$ or some approximation of $A$. This is considered a "model-free" paradigm in the following sense: the sources are estimated without any underlying assumptions about how they are created and mixed. There is no need to rely on a complex model with many unknown parameters for the system in which the signals are observed.

**The U1a sensor array and seismic data format**

An array of seismic sensors was installed at U1a and started collecting data continuously by April 2020. Within 4 months, over 2.7 TB of data was available. The sensors are INOVA Hawk model 3-axis geophones like that shown in Figure 3. The x and y axes are oriented with the horizontal/surface plane, and the z-axis corresponds to the "up/down" direction. A geophone essentially converts ground movement into a voltage reading. A spring-mounted wire coil in a magnetic field generates an electrical signal as it moves back and forth when subjected to forces from seismic or acoustic waves moving through the media in which it is deployed.

Figure 3: INOVA Hawk geophone

Initially, 36 sensors were arranged in roughly a "cross" pattern with 3 arms stretching out in approximately the N, W, and S cardinal directions, and the 4th arm diverging to the northeast. Each of these stations is identified with a letter corresponding to the branch, and a number, e.g. N01, N02, …, N13, S01, etc. In late 2020, another branch of arms was added to bring the total number of sensors up to 48. This branch was identified with the letter 'X'. Each sensor is buried under several inches of topsoil. An overhead view of the layout is shown in Figure 3.


Figure 4: Geophone layout

The sensors are time synchronized and able to sample at a rate of 500Hz, with a higher sampling rate possible for temporary periods or short-term experiments. Once downloaded onto a computer, the raw data is stored in what is called SEG-Y format, a file standard that was developed by the Society of Exploration Geophysicists. A single download from one sensor may be larger than 10 GB and represent 2 weeks of continuous recording. As most of the analysis and signal processing of this data was to be carried out using the Python scripting language, one of the first data engineering tasks was to be able to convert the raw seismic data from SEG-Y format to a numpy array. This required a parsing script which would comb through the byte stream of each SEG-Y file, skip

over irrelevant meta-data and file headers, and copy only the raw voltage readings from the x, y, and z axis. The only other data parsed from the file was the UTC time (down to 1 second resolution) of each digital record.

In addition to the SEG-Y parsing script, another script was developed to break each multi-day stream into usable numpy arrays representing 1 hr of data starting from the top of the hour. The script would sort each file into a directory ordered by year, month, and day, with each unique sensor ID belonging to the filename. With these "data management" utilities in place, the groundwork was laid out to allow easier searching, sorting, uploading, and analysis of the data.

**A low-rank linear decomposition approach to signal extraction**

This section explains one of the first methods devised for U1a signal extraction, based on a low-rank linear decomposition of cross-correlation aligned signals.

In many source separation settings, a simplifying assumption is that there are no time delays or time offsets with respect to a signal between different sensors. While this may be valid when recording speakers in a small room, as sound travels ~343 m/s in air, the assumption doesn't hold for seismic signals in the U1a array. The geophones are far enough apart that time delays between different sensors recording the same source(s) are evident. Thus, in order to construct our data matrix $X$ from Eq. (3), we do the following: assume there is some dominant signal which is evident in all of the array sensors (or perhaps a subset of them). When two signals are cross-correlated, the peak correlation index (or the argmax of the resulting correlation) gives the time offset of greatest similarity between two signals.

$$[X_i * X_j](n) \coloneqq \sum_{t=0}^{T-1} X_i(t) X_j(t + n \bmod T) \qquad (4)$$

Equation (4) gives the finite time discrete cross-correlation of two signals $X_i$ and $X_j$ of length $T$. By cross-correlating each pair of sensors, we can find the time offset between them with respect to the dominant signal and shift one so that the dominant signal is aligned. By doing this for all pairs, the dominant signal can be "built up" in a sense that for $M$ selected sensors, the $M$ rows of matrix $X$ are aligned in time. In other words, the no time delay assumption is artificially satisfied between sensors for the dominant signal.

Due to the shifted signals, the cross-correlated data $X$ will have a slightly shorter length in time, but this is irrelevant, as the overall signal duration is assumed to be much longer than the dominant signal in question. The linear decomposition approach takes aligned matrix $X$ and approximates it with matrices $W$ and $H$:

$$X \approx W \cdot H \qquad (5)$$

$W$ and $H$ are typically of lower rank, in order to achieve data reduction.

For example, assume signal is made up of a linear combination of two main components: $S_1(t)$ and $S_2(t)$. At time $t$, the signal from sensor $X_i$ can be decomposed as

$$X_i(t) = a_{i1} S_1(t) + a_{i2} S_2(t) + \epsilon_i(t) \qquad (6)$$

With the error term accounting for all other aspects of the signal not from the two main components. Equation (6) expresses this decomposition as a rank 2 approximation of $X$. The $i$th row of $W$ is $[a_{i1} \;\; a_{i2}]$ and the two columns of $H$ are the sources $S_1(t)$ and $S_2(t)$. A schematic is shown below in Figure 5.

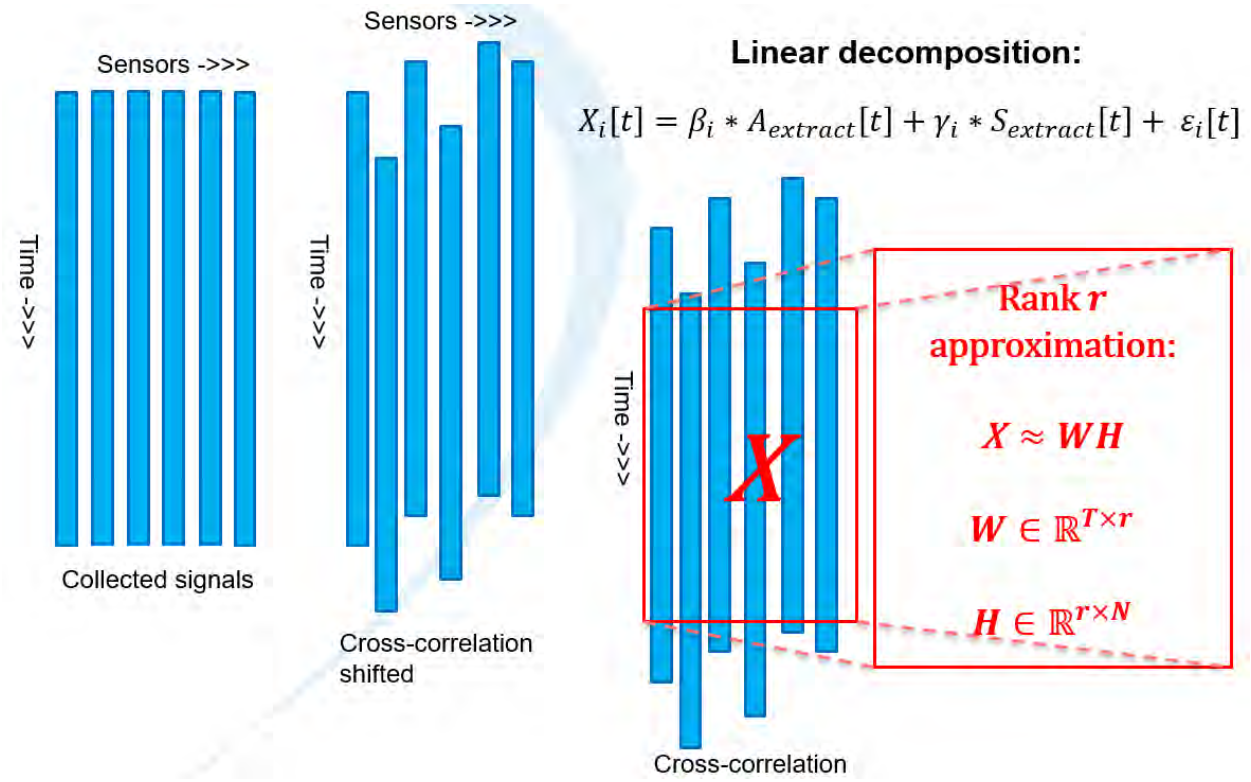


Figure 5: Cross-correlated and aligned signals make up signal matrix $X$.

Thus, the task of extracting two signals from matrix $X$ comes down to finding the matrices $W$ and $H$ in the factorization in Equation (5). This can be achieved in several ways, two common approaches are the principal component analysis, and non-negative matrix factorization, which can be expressed via the optimization:

$$\textbf{PCA: } \underset{W \in O(n)}{\operatorname{argmin}} \ \|X - W \cdot H\|_F^2 \qquad (7)$$

$$\textbf{NMF: } \underset{W,H \geq 0}{\operatorname{argmin}} \ \|X - W \cdot H\|_F^2 \qquad (8)$$

where $O(n)$ is the space of orthogonal $n \times n$ matrices and $\|\cdot\|_F$ is the Frobenius matrix norm. The PCA solution is efficiently computed via singular value decomposition. Non-negative matrix factorization requires that $X, W$, and $H$ all have non-negative entries and will be discussed in the next section.

After computing the PCA and obtaining $W$ and $H$, we may approximate $X$ with a rank-1 matrix formed by taking the outer product of the first principal component. From here, find the residual matrix $R = X - X'$, "unshift" $R$ using the time offsets determined by the sensor cross-correlations, then repeat the process.

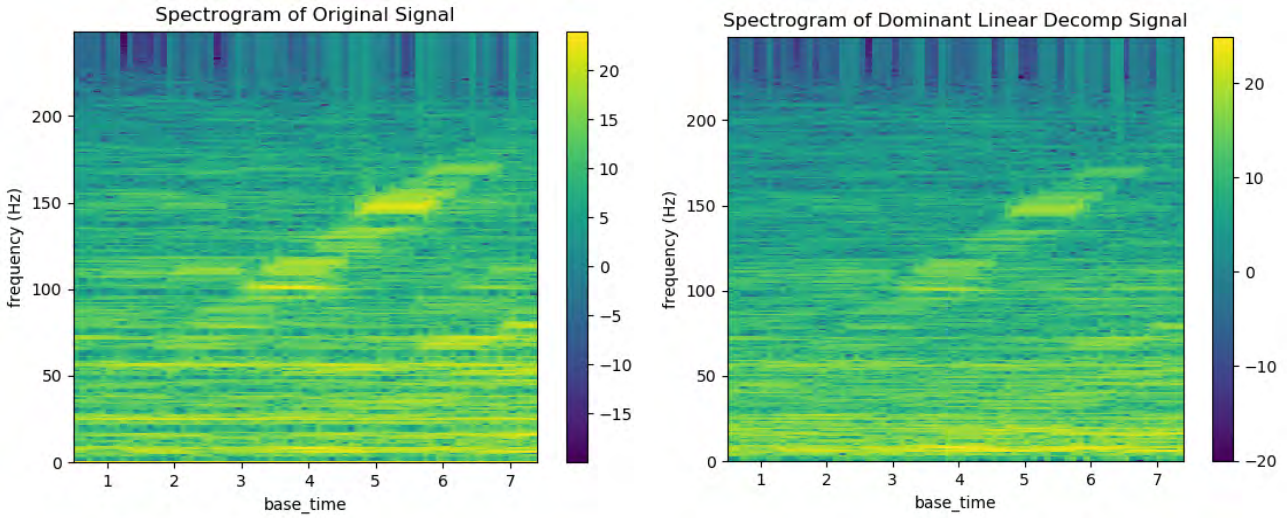


Figure 6: Spectrograms of original and correlated signal.

This method was applied using five sensors in an attempt to develop an analytic signature for the U1a elevator shaft motor, characterized by a sloping bright pattern in the spectrogram. As Table 1 shows, 3 principal components were enough to describe much of the variance in the two North arm sensors. Different types of signals and known patterns have yet to be analyzed with this technique, but it is a good start to a low rank signal decomposition method with the seismic data.

| Proportion of original signal variance variance explained with linear decomp model | | | | | |
|---|---|---|---|---|---|
| | **N05** | **N13** | **S03** | **W03** | **E03** |
| **1 component** | 0.2031 | 0.1235 | 0.4895 | 0.2911 | 0.2799 |
| **2 components** | 0.4628 | 0.5365 | 0.4953 | 0.4805 | 0.5476 |
| **3 components** | 0.9549 | 0.9003 | 0.5029 | 0.4840 | 0.5481 |
| **4 components** | 0.9606 | 0.9059 | 0.5807 | 0.9999 | 0.7762 |

Table 1: PCA variance results

**Monaural signal separation with non-negative matrix factorization**

This section describes another commonly used method for blind source separation: non-negative matrix factorization (NMF). There has been a considerable amount of research and literature produced on NMF since the late 1990's, due to its usefulness in high-dimension data modeling and machine learning applications [source]. A few applications include document topic analysis, auditory scene analysis, electroencephalography, image and video separation, and recommender systems.

Like with the PCA method, we assume the data $X$ can be approximated with a linear decomposition: $X \approx W \cdot H$. As the name implies, there is a constraint that all three matrices have non-negative values. This is useful because it does not allow any parts of the sources to cancel out. The intuition here is that NMF is a "parts-based" representation of the signal $X$.

Finding $W$ and $H$ can be posed as an optimization problem:

$$\{W, H\} = \underset{W,H}{\operatorname{argmin}}\ D(X \,||\, W \cdot H)$$

$$\text{Subject to: } W, H \geq 0 \qquad (9)$$

Where $D(\cdot \,||\, \cdot)$ is a cost function, or divergence (e.g. Kullback-Leibler). A further constraint on the rank of $W$ and $H$ can be imposed. The solution $\{W, H\}$ is not unique in the sense that an equivalent factorization can be found with a different scaling and permutation (order) of the basis elements of $W$ and $H$. Currently, no algorithms have been shown to converge on a global-optimal solution. However, local optima have shown to be useful in practice.

I will provide here an overview of the main algorithms and techniques which are used to find a non-negative factorization of $X$. The dissertation [MS] gives a succinct overview of NMF with original sources. A nice geometric intuition of NMF is explained in [AA]. If the data $X$ are non-negative, each observation lies inside the positive orthant of the corresponding feature space, that is, the columns of $X$ lie within a convex cone (cf Fig. 7). By normalizing the columns of $H$, the data is mapped to the convex hull with vertices represented by the columns of $W$ (Fig. 8).

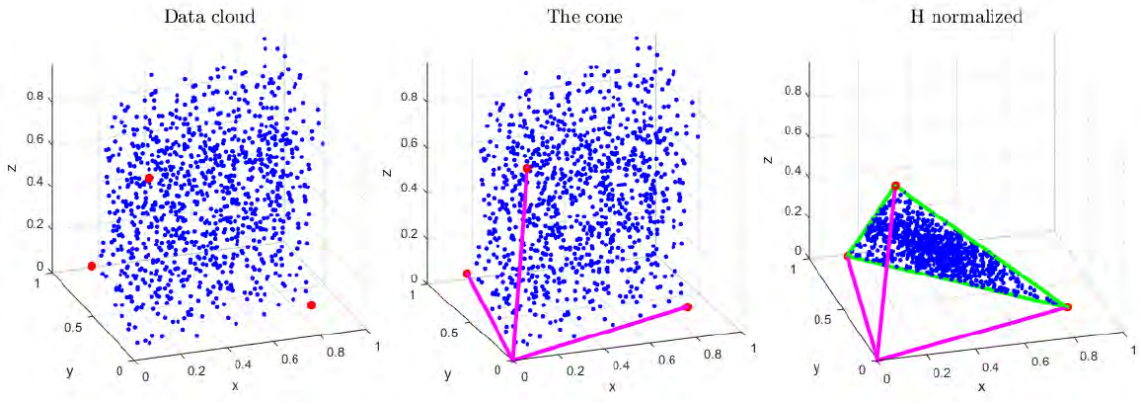


Figure 7: NMF as a convex cone [AA]

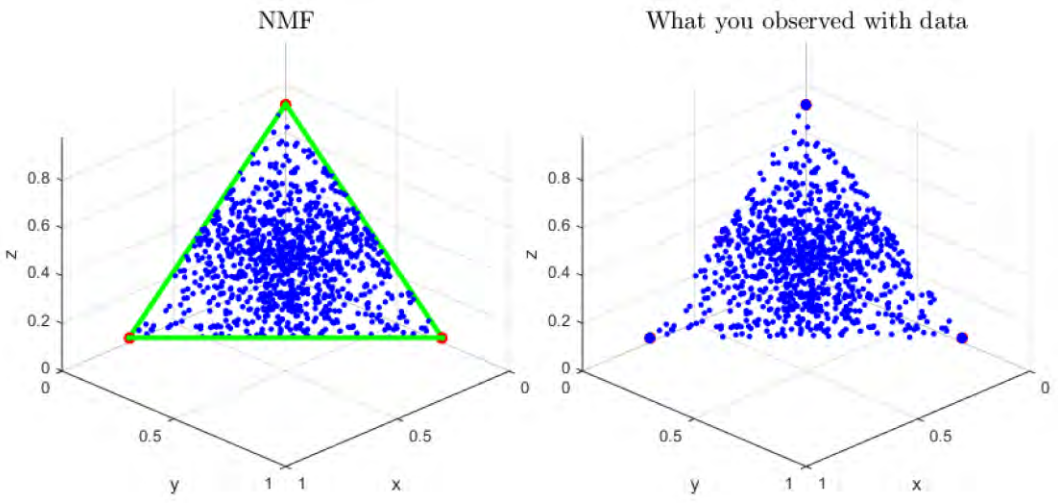


Figure 8: Finding $H$ amounts to finding the vertices of the triangle [AA]

The optimization methods for NMF fall into 3 different types: direct optimization, alternating optimization, and alternating descent [MS]. Direct optimization of Eq. (9) requires solving the non-negative constrained, non-linear problem but can be infeasible due to the high number of parameters. Alternating optimization methods treat the search for $W$ and $H$ as two separate subproblems. Each iteration to solve for one matrix is taken while the current step for the other matrix is held fixed, until both converge to a local solution for the main problem.

Alternating descent methods are similar to alternating optimization, except at each step, the cost/divergence function for each separate matrix is only reduced, not minimized. An example of this type of algorithm is the multiplicative update rule given by Lee and Seung [DL]. The authors provide two theorems, which state that the Frobenius norm (Euclidean distance) and KL divergence is non-increasing under the update rules. The rule for the Euclidean distance is given here.

**Theorem:** The Euclidean distance $\|X - WH\|$ is nonincreasing under the update rules

$$H_{\alpha\mu} \leftarrow H_{\alpha\mu} \frac{(W^\top X)_{\alpha\mu}}{(W^\top WH)_{\alpha\mu}}$$

$$W_{ia} \leftarrow W_{ia} \frac{(XH^\top)_{ia}}{(WHH^\top)_{ia}}$$

The Euclidean distance is invariant under these updates if and only if $W$ and $H$ are at a stationary point of the distance.

This is guaranteed to converge to a locally optimal solution, which solves the problem from Eq. (9). The proof of this update rule is done with an auxiliary function. The ease of computational implementation has made Lee and Seung's update rule algorithm one of the most widely used for NMF [MS].

We now turn to the application of NMF methods on the seismic signal separation task. As explained earlier, the raw data from the geophones represent voltage values which fluctuate as seismic and acoustic energy waves vibrate the ground around the sensor. While NMF appeared to be a promising tool in our signal separation task, the voltage signals in the seismic data are not inherently non-negative, as they are in other applications. A short sample of the raw voltage signal is shown in Fig 9.

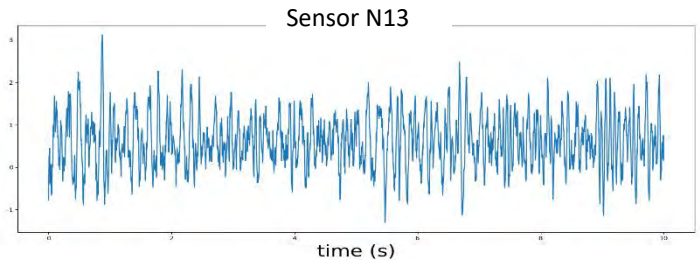


Figure 9: Raw voltage signal from seismic sensor.

Some naïve work-arounds were initially considered: (1) the data could be "min-shifted," that is, add the minimum value to the signal, but this would clearly misrepresent a "loud" seismic event (with high amplitude voltage readings) because the maximum negative value would be close to zero and equate to an absence of signal. (2) the absolute value of the voltage could be taken, but here again, values of zero would be present where there is high amplitude oscillations.

A method to separate the sources from one sensor, known as monaural separation, is described in [NB]. The technique used is to represent the single microphone recording as a spectrogram, by applying the short-time Fourier transform (STFT). As applied to the U1a seismic array, this analysis focuses only on separating sources within a single sensor, so the time offset and cross-correlation preprocessing in the previous section doesn't apply. The STFT of the signal from one geophone turns the one-dimensional time series voltage reading into a matrix of positive values: the rows represent each discretized spectra or frequency, the columns represent the sampled times, and the values are the weight of each frequency occurring at the particular time.

The discrete STFT is defined as follows [NB]:

$$X_m(\omega_k) = e^{-j\omega_k mR} \sum_{t=-N/2}^{N/2-1} x(t+mR)w(t)e^{-j\omega_k t}$$

$x(t) =$ input signal at time $t$

$w(t) =$ length $M$ window function (e.g. Hann, zero-centered Gaussian)

$N =$ DFT size, in samples

$R =$ hop size (samples) between successive DFT

$M =$ window size, in samples

$\omega_k = 2\pi k/N$ , $k = 0,1,2,\dots,N-1$

Note that this is just the Fourier transform of the signal combined with a window function $w(t)$. The result $X_m(\omega_k)$ is a complex-valued matrix which represents the magnitude and phase of the signal in each point in time and frequency. The spectrogram is obtained by taking the magnitude squared: $|X_m(\omega_k)|^2$ .

This representation of the data can now be input for one of the non-negative matrix factorization algorithms discussed. In the case of an audio recording sample, consider the interpretation of the columns and rows of the matrices $W$ and $H$, respectively. Suppose we have a rank 3 approximation of the spectrogram $X$. The $i$th column of $X$ is the matrix-vector product $Wh_i$, where $h_i$ is the $i^{th}$ column of $H$. In turn, $Wh_i$ is a linear combination of the columns of $W$, scaled by the elements in $h_i$. Since a column of $X$ represents the amount of each spectra present in the signal at time $t$, the interpretation is that each column of $W$ represents a "part" or "source" in the signal $X$, and the columns of $H$ represent the "activations" or "gains" of each source at time $t$.

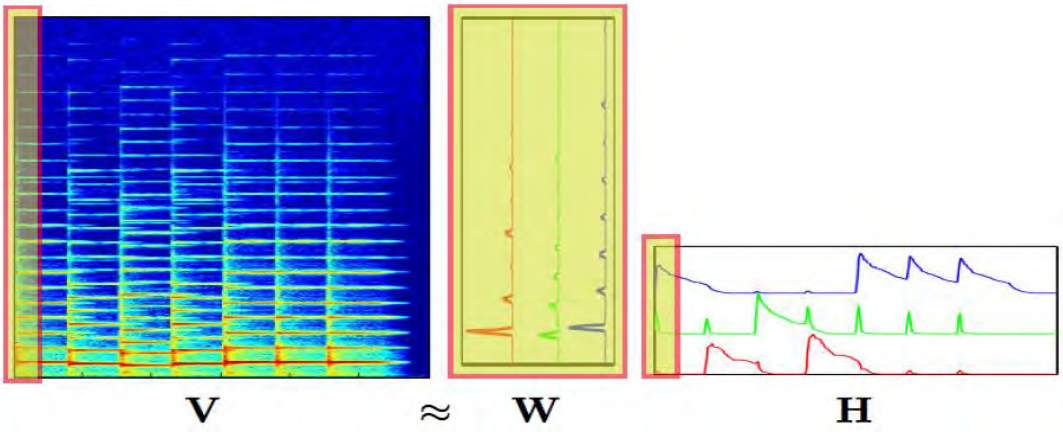


Figure 10: NMF of an audio spectrogram [NB]

The audio spectrogram in Fig. 10 shows this interpretation. For a rank 3 decomposition, the columns of $W$ are the frequency domain representation of the sources, while the rows of $H$ are the gains for each source. Equivalently, the spectrogram $V$ can be thought of as the sum of three matrices formed by taking the outer product of each column of $W$ and row of $H$.

In order to reconstruct each separated signal component (i.e. each column of $W$) in the time domain, the outer product of corresponding columns of $W$ and rows of $H$ is computed. This is the spectrogram of the separated component. This, along with phase info from the original STFT, makes up the input for the inverse short-time Fourier transform (ISTFT), which reconstructs the original time domain signal [NB]. A consideration when using the STFT is the type of window function $w(t)$ to select. One common window function $w(t)$ is the Hann window, which is essentially one period of a squared cosine wave. It is the default window in most software and the one applied in our use case.

A demonstration of this workflow is now shown. Consider the standardized signal $X(t)$ shown in Fig. 11. It is an unfiltered 2 min 40 sec recording from one axis of geophone station N13. We will apply NMF using three components to attempt signal separation.

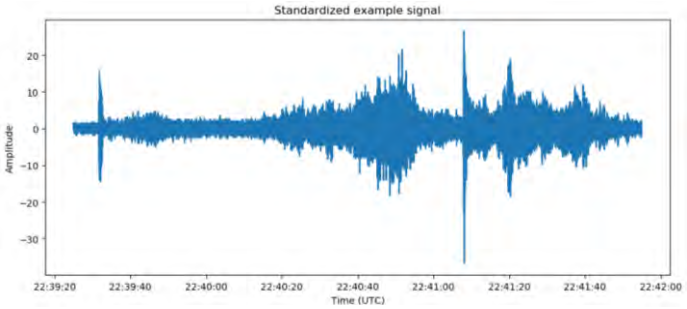


Figure 11: Standardized voltage signal, N13

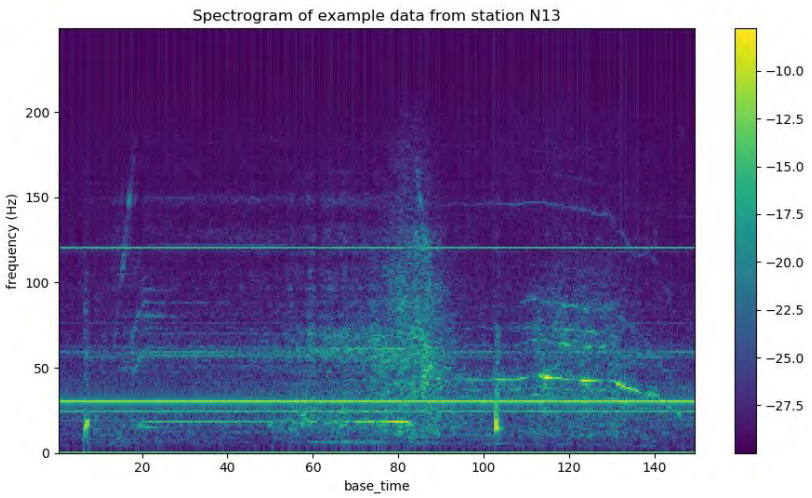


Figure 12: Spectrogram for working example

We can see from the spectrogram that several things are going on. There are obvious constant nuisance tones around 25-30 Hz and ~120 Hz. We see the elevator shaft signature in the form of the steep diagonal positive slope line followed by a plateau, and part of the diagonal negative slope line being obscured by a broad spectrum signal from around 60-90 seconds. There appear to be high amplitude, short duration "bursts" around 7 seconds in and 105 seconds in. From actual observations of elevator operation, these features are most likely the elevator door sliding open and shut, which creates a loud bang. There is also a varying, narrow spectrum tone which gradually drops off by the end of the recording.

After applying the STFT to this signal with a Hann window of length 256, the squared modulus of the complex values is used to make our input matrix $X$ for the NMF. The default cost (or loss) function used is the squared Frobenius norm:

$$\|X\|_F^2 = \sum_i \sum_j |X_{ij}|^2$$

The matrix decomposition software returns $W$ and $H$, computed using the multiplicative update rule shown earlier. In order to reconstruct the 3 source estimates, let $\widehat{X}_i = W_i H_i^{\mathsf{T}}$ be the estimated modulus for the STFT of the $i$th source, where $H_i^{\mathsf{T}}$ is the $i$th row of $H$. This is multiplied by the phase values for the STFT of $X$ and given as input to the inverse STFT function. Our estimate source signal $\hat{x}_i(t)$ is returned.

The result of this workflow on the example signal is shown in Fig 13. We have the three source estimates plotted in the time domain. They are slightly shorter due to the sliding window used in the STFT, which is ok because the signals of interest lie well within the timeframe of our sample. We can see that NMF has produced distinct estimates. Source 1 appears to have low amplitude background noise until around 22:40:50, before a signal appears. The Source 2 plot is scaled vertically in amplitude but shows a different signature than source 1. Most activity is from 22:40:20 to 22:41:00. The third source clearly

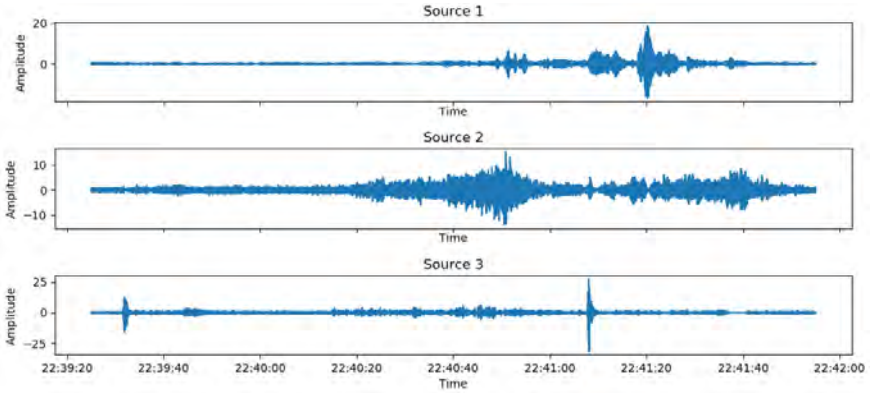


Figure 13: Recovered source estimates from rank-3 NMF

shows a separation of the loud “bang” noises coming from the elevator shaft door, occurring at 22:39:30 and again at 22:41:07. Although the amplitude scaling on this plot is higher, we can see that the rest of source three appears to have less activity between the impulse signatures.

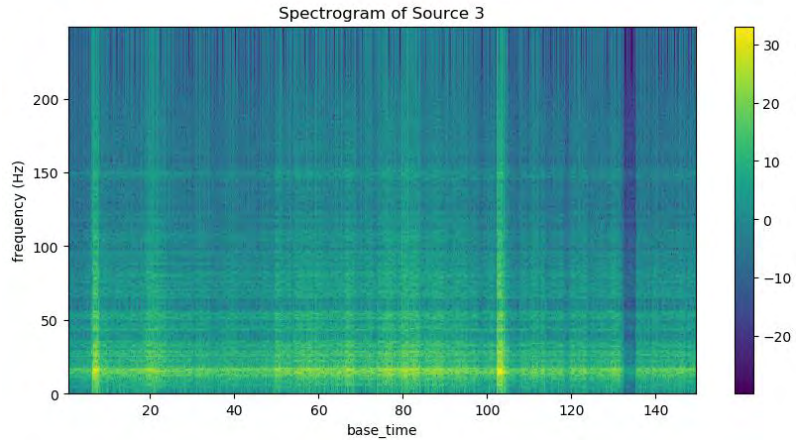


Figure 14: Recovered source 3 spectrogram (impulse signatures)

What does a spectrogram of one of these sources look like? In Fig 14, the Source 3 spectrogram is shown. Note the color scaling is on a much broader range, from -20 to 30, compared to the scale from -27 to -10 on the unseparated signal spectrogram. We can clearly see the two impulse sources caused by the slamming shaft door, represented as two bright vertical bars.

There remains much to be investigated when it comes to applying NMF to the seismic signal separation. Different algorithm parameters and loss functions, or divergence measures, can be experimented with. The Kullback-Leibler and Itakura-Saito divergences are two examples.

The workflow laid out above only attempts monaural separation from a single recording. There are possibly other non-negative representations of the seismic data. One such method could come from applying the Hilbert transform to the $i$th sensor signal $X_i(t)$. This obtains its analytic representation, and the instantaneous amplitude is a positive signal representing the “envelope.” The envelope signals from multiple sensors could then be aligned in a matrix $X$ and decomposed as a lower rank factorization. Further investigation would be needed to test the concept of this method.

**Ground truth collection experiment at U1a**

In July 2021, myself and another student intern, UNLV mathematics PhD candidate Eric Jameson, had the opportunity to visit the U1a site and actually create various short duration, low energy signals in the proximity of one of the Hawk geophones. The goal of this experiment was to make a ground truth collection so that our various analytics could be measured for quality of signal separation. This would hopefully make it easier to test versions of signal separation methods, and see which ones would be most useful.

The experiment was conducted rather ad hoc, as we had no prior experience or knowledge with subjecting the geophones to artificial “sources.” The main idea was to create seismic and acoustic signals which would for certain be picked up by station X12, then record the time at which the various signals were induced, and download the data afterwards. This would provide us with a labeled data set which could be used to demonstrate how well different methods work. It could also serve as a prototype for a future supervised learning method training set.

This section will discuss the result of applying NMF to our collected data using two of the artificial source signals we created. The sources were (1) a government vehicle driving slowly (<15 kph) around the sensor in

approximately a circle, and (2) a 2-3 cm diameter steel ball bearing being dropped from about 1 meter high onto a thick steel plate pushed firmly into the ground.

The Hawk geophone station which we employed for this experiment was station X12, located within 50 meters of relevant site activity including the U1a elevator shaft, a ventilation fan, and a passing utility road. The surrounding area was mostly desert topsoil and dry grass. The entire experiment spanned roughly 1 hours between 11:30 am and 12:30 pm PDT. This time included creating sources other than the two explained below.

The procedure started with recording each source individually. First, the government vehicle was driven for 2 laps around the sensor in roughly a circle, with each lap taking approximately 90 seconds. Next, the steel plate was placed within 1 meter of the geophone and pressed into the ground by standing on it. The record keeper and stopwatch operator would then give the signal for the ball bearing dropper to release the steel ball and let it hit the plate. This was repeated twice, for each distance of 1, 4, 8, and 12 meters from the geophone.

With the sources recorded in isolation, we next allowed the vehicle to be driven while repeating the ball bearing drop procedure, so both sources would be simultaneously picked up by the sensor. The unfiltered signal for this window of time is shown in the spectrogram in Fig 15. The broad-spectrum feature comes from the vehicle, and the sharp vertical lines are from the ball bearing striking the plate. The monaural NMF method explained was used to decompose this data into two sources, the rational being that one source would be the vehicle, and the other would be the ball bearing.

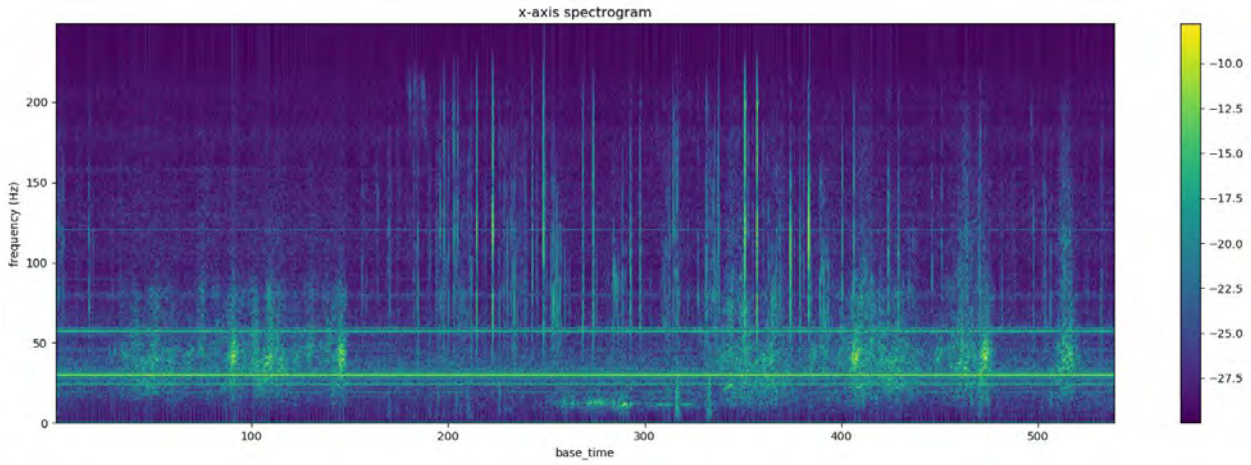


Figure 15: Geophone X12 spectrogram of ball bearing and vehicle signature overlap.

The results of the NMF workflow are shown in the plots of Figure 16. The top plot is the standardized and bandstop filtered original signal. The bandstop filtering was done to attenuate the constant nuisance tones clearly visible in the spectrogram, at around 25-30 Hz and 60 Hz. The bottom plot is an overlay of two “label” or indicator vectors for the presence of each source. That is, if the vehicle was driving at a certain time, or the ball bearing was dropped, the vector element at that time index would contain a 1, and 0 if the source was not present. The NMF estimates of sources 1 and 2 are plotted on the second and third rows. The vertical axis scaling of the plots for source 1 and 2 are different because the amplitude of the ball bearing impact was so much higher than the broad-spectrum source coming from the vehicle. It is easy to see from the plots that Source 2 is the separation of the ball bearing source, and Source 1 contains most of the signal from the car.

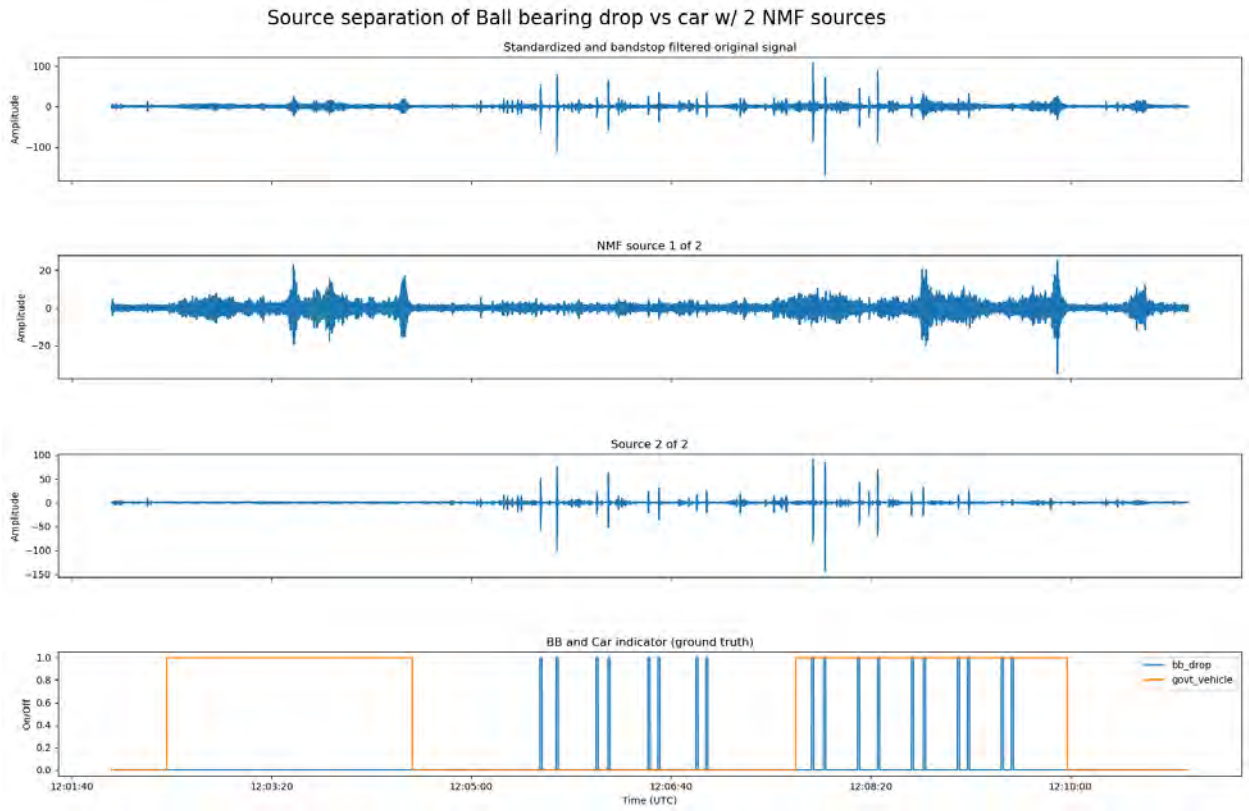


Figure 16: Rank-2 NMF separation results for ball bearing and vehicle experiment.

Figure 17 shows a zoomed-in version of the four plots. Note the time window from 12:07:40 to 12:08:30. Source 1 is almost completely absent of the high amplitude signature coming from the ball bearing. These results are a promising step toward the development of a fully unsupervised source separation capability for the U1a seismic array. As with the example in the NMF workflow discussion, NMF appears to excel in separating short duration, high amplitude sources from the background activity. A possible use case would be to apply the monaural method to each seismic station and extract the short duration impulse events, and use the separated impulse signals for pairwise sensor cross-correlation.

This one experiment was our only chance over the summer to collect ground truth, but it provided a valuable data set as well as laid the groundwork for future experiments.

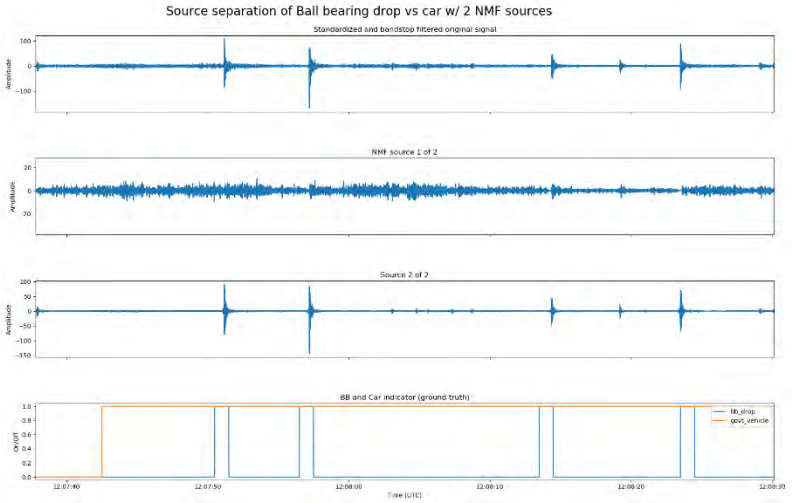


Figure 17: Closer view of separation results.

**LANL Research Article**

Before concluding this report, it is worth discussing the most interesting publication [FI] I encountered during the course of literature review for NMF applications. It is an article by four authors from Los Alamos National Labs, titled "Nonnegative Matrix Factorization for identification of unknown number of sources emitting delayed signals." They present a method in which NMF is used to identify an unknown number of mixed signals, delays between sensors, propagation speed of the signals, and location(s) of the sources, using a sensor array. Discovering this article was a nice surprise, as it considers essentially the exact same objective of the U1a seismic separation project! It even mentions nuclear nonproliferation as a possible application. An algorithm called "ShiftNMF*k*" is introduced, which is capable of determining the "unknown number of the sources of delayed signals and estimating their delays, locations, and speeds based only on records of their mixtures."

An overview of this algorithm is as follows: first, a version of the NMF optimization which accounts for shifted signals is modified to stop at an optimal number of iterations. Then, a system to discard solutions when the optimization is performed many times is devised, followed by a clustering method used to determine the actual number of sources. A further optimization procedure estimates the locations of the sources, and finally a Bayesian uncertainty analysis is done to determine how accurate the locations are. The authors provide a link to the Git repository which contains the code used to implement this algorithm. It is implemented in the Julia language. Given more time to delve into this paper, my top priority would be to attempt applying their algorithm to work on our seismic array data. Their results are very interesting and likely the best currently developed method for the U1a source separation objectives.

# CONCLUSIONS AND FURTHER RESEARCH

Despite the promising initial results of NMF on our data, it was very challenging to actually devise a scheme to quantitatively measure the quality of extracted signals. It is easy to visually confirm that monaural NMF is able to indicate the presence of an impulsive source, as in the ball bearing experiment. But without more labeled data, it's difficult to say what the next step should be. There still are many possible questions and topics worth investigating, such as: can we separate moving sources from stationary sources? Can multilateration results be improved by using NMF? Can we create artificial sources which are picked up by multiple sensors, rather than small localized sources? Do the surface acoustic sources differ in some fundamental way from the subterranean seismic sources?

There are also a myriad of data pre-processing and transformation tasks to consider, such as different correlation and signal stacking techniques, application of classical signal processing filters, and better data engineering to make it easier to select and manipulate data of interest. Visualization and animation tools can prove extremely useful for understanding the types of signals we are trying to analyze. In addition to this, the data collection at U1a spans different time resolutions and modalities, as mentioned in the introduction. One topic of interest is the combination of different data streams to develop some type of combined analytic. This could be applied in signal extraction for example if power usage data was able to provide some amount of prior knowledge for the blind separation problem.

Overall, there is still much room for analytic development regarding the source separation task at U1a. It is my hope that these initial explorations of applying NMF to the seismic data can be continued and inspire similar or perhaps completely different methods to be attempted, whether it is for the ADAPD venture or other nonproliferation projects.

# REFERENCES

[AA] Andersen Ang (2018) “What’s happening in Nonnegative Matrix Factorization? Models and Algorithms.” Structured Low-Rank Matrix/Tensor Approximation. Leuven, Belgium. Accessed at https://angms.science/talks.html.

[DL] Daniel Lee & H. Sebastian Seung (2000) “Algorithms for Non-negative Matrix Factorization.” Advances in Neural Information Processing Systems 13. NIPS 2000. Accessed at https://papers.nips.cc/paper/2000/hash/f9d1152547c0bde01830b7e8bd60024c-Abstract.html.

[FI] Iliev FL, Stanev VG, Vesselinov VV, Alexandrov BS (2018) “Nonnegative Matrix Factorization for identification of unknown number of sources emitting delayed signals.” PLoS ONE 13 (3): e0193974. https://doi.org/10.1371/journal.pone.0193974

[MS] Schmidt, M. N. (2009). “Single-channel source separation using non-negative matrix factorization.” Technical University of Denmark, DTU Informatics, Building 321. IMM-PHD-2008-205.

[NB] Nicholas Bryan & Dennis Sun (2013). “Source separation tutorial mini-series II: introduction to non-negative matrix factorization.” Center for Computer Research in Music and Acoustics, Stanford. Accessed at https://ccrma.stanford.edu/events/source-separation-tutorial-mini-series-ii-introduction-non-negative-matrix-factorization .