Principal Component Analysis
And its application to Relativistic Heavy-Ion Collisions

For HENPIC
Ziming Liu
Peking University
May 30, 2019
Overview

• Machine Learning Overview
• Machine Learning in Physics
• Principal Component Analysis in Heavy-Ion
I. Machine Learning
Machine Learning in this world
Categories of ML

- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning
Example of DL: dogs vs. cats

Supervised learning
experience a dataset contains many features, and each example is also associated with a label or target.

Example of DL: dogs vs. cats

Unsupervised learning
experience a dataset contains many features **but without labels**, and learn useful properties of the structure of this dataset.

Reinforcement learning

concern with how software agents ought to take actions in an environment so as to maximize some notion of cumulative reward.

WIKI: https://en.wikipedia.org/wiki/Reinforcement_learning

David Silver et al., Nature 529(2016) 484-489
Computer Vision

Image style transition


Image generation

A. van den Oord et al., NIPS, (2016), arXiv: 1606.05328
Language Processing

Machine translation

<table>
<thead>
<tr>
<th>Language</th>
<th>Original Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multi src</td>
<td>Bei der Metropolitního výboru pro dopravu für das Gebiet der San Francisco Bay erklärt Beamte, der Kongress könne das Problem bankrottstvo доверительного Фонда строительства шоссейных дорог einfach durch Erhöhung der Kraftstoffsteuer lösen.</td>
</tr>
<tr>
<td>EN ref</td>
<td>At the Metropolitan Transportation Commission in the San Francisco Bay Area, officials say Congress could very simply deal with the bankrupt Highway Trust Fund by raising gas taxes.</td>
</tr>
<tr>
<td>bpe2char</td>
<td>During the Metropolitan Committee on Transport for San Francisco Bay, officials declared that Congress could solve the problem of bankruptcy by increasing the fuel tax bankrupt.</td>
</tr>
<tr>
<td>char2char</td>
<td>At the Metropolitan Committee on Transport for the territory of San Francisco Bay, officials explained that the Congress could simply solve the problem of the bankruptcy of the Road Construction Fund by increasing the fuel tax.</td>
</tr>
</tbody>
</table>

Chinese poetry generation

Speech recognition
W. Xiong et al., IEEE/ACM Transactions on Audio Speech & Language Processing, 2016, PP(99), arXiv: 1610.05256
What is a neural network?
Define Functions

Neural Network

input  hidden layer  output

"dog"

"cat"
Neuron

\[ y = \sigma(z), \quad z = w_1 x_1 + w_2 x_2 + \ldots + w_n x_n + b \]

\( \sigma \) is a simple non-linear activation function

\( w_1, w_2, \ldots, w_n, b \) is the trainable parameters
Neural Network Structures

M. Ranzato, NEURAL NETS FOR VISION,
Gallery of neural network models
II. Machine Learning in Physics
A win-win game

Why does deep and cheap learning work so well?

arXiv:1608.08225
Physics Machine learning

https://physicsml.github.io/pages/papers.html

PAPERS

The following are recent papers combining the fields of physics - especially quantum mechanics - and machine learning. Please email Anna Go if you would like to see a paper added to this page.

APPLYING MACHINE LEARNING TO PHYSICS


- "Optimizing Quantum Error Correction Codes with Reinforcement Learning", Hendrik Poulsen Nautrup, Nicolas Delfosse, Vedran Dunjko, Hans J. Briegel, Nicolai Friis, arXiv: 1812.08451, 12/2018

- "Parameters optimization and real-time calibration of Measurement-Device-Independent Quantum Key Distribution Network based on Back Propagation Artificial Neural Network", Feng Yu Lu, Zhen Qiang Yin, Chao Wang, Chao Han Cui, Jun Teng, Shuang Wang, Wei Chen, Wei Huang, Bing Jie Xu, Guang Can Guo, Zheng Fu Han, arXiv: 1812.08388, 12/2018


Machine learning in physics

• Condensed matter physics
• Astrophysics
• High Energy Physics
• Relativistic Heavy-Ion physics
Machine learning in physics

- Condensed matter physics
  - Ising model: Phase Transition
  - Renormalization group
  - Faster sampling than MCMC
- Astrophysics
- High Energy Physics
- Relativistic Heavy-Ion physics
Ising model

Juan Carrasquilla, Roger G. Melko

Feed-forward neural network

T-SNE visualization for high and low temperature configuration

Predict transition temperature

Machine learning can find phase transition from data without any guidance!
Ising model

arXiv:1410.3831
Pankaj Mehta, David J. Schwab

Ising model can be calculated efficiently with deep learning!

Ising model ⇔ Deep learning
Renormalization group ⇔ Pooling operation

Pooling size = 2

Length = 6
Length = 3
Ising model
Phys. Rev. Lett. 122, 080602
Dian Wu, Lei Wang, and Pan Zhang

VAN: Sampling a configuration from a neural network
VAN outperforms traditional methods in accuracy

Deep generative models can generate samples better than MCMC!
Machine learning in physics

• Condensed matter physics
• Astrophysics
  • Gravitational lensing
  • Blackhole
• High Energy Physics
• Relativistic Heavy-Ion physics
Parameter estimation with Neural network
10^7 times faster than traditional method

Input images of gravitational lensing

Machine learning helps automated analysis of parameters of gravitational lensing!
Take a photo of the blackhole

Reconstructing blackhole with missing data

Machine learning can help reconstruct corrupted images.
Machine learning in physics

• Condensed matter physics
• Astrophysics
• High Energy Physics
  • Higgs identification
  • Classifying Quark/Gluon jet
  • Generate jet images
• Relativistic Heavy-Ion physics
Deep learning can help search for exotic particles and increase the accuracy 8% for traditional methods.
Quark/Gluon jet tagging

JHEP 1701 (2017) 110
Patrick T. Komiske, Eric M. Metodiev (MIT, Cambridge, CTP), Matthew D. Schwartz

Deep learning can outperform traditional methods in discriminating between quark jet and gluon jet
Generate jet images

PhysRevD.97.014021
Michela Paganini, Luke de Oliveira, and Benjamin Nachman

Three-dimensional representation of a 10 GeV $e^+$ incident perpendicular to the center of the detector.

GAN: generative adversarial network

discriminator

Deep Learning for generating 3D High Energy Particle Showers Events!
Machine learning in physics

- Condensed matter physics
- Astrophysics
- High Energy Physics
- Relativistic Heavy-Ion physics
  - Identify QCD EOS
  - Lattice
  - pp/AA jet discrimination
  - Deep learning hydrodynamics
Heavy-Ion: EOS of QGP

Nature Communications volume 9, Article number: 210 (2018)
Long-Gang Pang, Kai Zhou, Nan Su, Hannah Petersen, Horst Stöcker & Xin-Nian Wang

Phase diagram of QCD

Convolutional neural network
Heavy-Ion: EOS of QGP

Nature Communications volume 9, Article number: 210 (2018)
Long-Gang Pang, Kai Zhou, Nan Su, Hannah Petersen, Horst Stöcker & Xin-Nian Wang

What does Neural Network learn?

Prediction accuracy

Machine learning can distinguish QCD phase!
Lattice

arXiv:1810.12879
Kai Zhou, Gergely Endrődi, Long-Gang Pang, Horst Stöcker

Neural network architecture

Prediction of density and the squared field nicely detect the critical point

Regressive and generative neural networks for complex scalar field theory!
Jet in Heavy-Ion

Yang-Ting Chien

gluon jets broader distributions compared to quark jets

JEWEL+PYTHIA [2.76 TeV]

medium interactions $\rightarrow$ broader energy distribution for both quark and gluon jets

Train a CNN to distinguish

$\begin{cases}
\text{Quark jet} & \{ \text{pp} \} \\
\text{Gluon jet} & \{ \text{AA} \}
\end{cases}$

Neural network to benchmark classification performances!
Heavy-Ion: DL of Hydro

arXiv: 1801.03334
Hengfeng Huang, Bowen Xiao, Huixin Xiong, Zeming Wu, Yadong Mu, Huichao Song

\[ \partial_\mu T^{\mu \nu} = 0 \]
Heavy-Ion: DL of Hydro

arXiv: 1801.03334
Hengfeng Huang, Bowen Xiao, Huixin Xiong, Zeming Wu, Yadong Mu, Huichao Song
Heavy-Ion: DL of Hydro

arXiv: 1801.03334
Hengfeng Huang, Bowen Xiao, Huixin Xiong, Zeming Wu, Yadong Mu, Huichao Song

Machine learning can mimic hydrodynamical evolution!
III. Principal Component Analysis (PCA) in Heavy-ion Physics

arXiv: 1903.09833
Ziming Liu, Wenbin Zhao, Huichao Song
Principal Component Analysis (PCA) of Collective Modes in Relativistic Heavy-Ion Collisions

Recent work, paper in preparation
Ziming Liu, Arabinda Behera, Huichao Song, and Jiangyong Jia
Reconsideration on studying sub-leading flow with PCA
Principal component analysis of event-by-event fluctuations

PRL 114, 152301 (2015)

Rajeev S. Bhalerao, Jean-Yves Ollitrault, Subrata Pal, and Derek Teaney

Non-zero sub-leading flow?
What is PCA?
An intuitive way for PCA

Eigenvalues $\sigma$: importance of eigenvectors

Eigenvectors $z$: correlations between features
Principal component analysis — Application

Dataset: Faces of different people

Eigenfaces
Principal component analysis —— Application

- Each face is decomposed into superposition of eigenfaces.

\[
\hat{x} = \mu + w_1 u_1 + w_2 u_2 + w_3 u_3 + w_4 u_4 + \ldots
\]

- We can drop unimportant high order eigenfaces. So we can use a few coefficients and corresponding eigenfaces to reconstruct original faces. (Image compression)

\[
\begin{align*}
P &= 4 \\
P &= 200 \\
P &= 400
\end{align*}
\]
Principal component analysis --math

**Theorem: SVD (Singular Value Decomposition)**
For a complex (real) matrix $A \in \mathbb{R}^{n \times m}$, \exists unitary (orthogonal) matrices $U_{n \times n}$ and $V_{m \times m}$, along with a sub-diagonal matrix $\Sigma_{n \times m}$ such that
$$A = U \Sigma V$$

Where $\Sigma = diag(\sigma_1, \sigma_2, \ldots,)$ such that $\sigma_1 \geq \sigma_2 \geq \ldots \geq 0$

$\sigma$: singular values
$v_i$: eigenmodes/ eigenvectors/principal components

![Diagram](image.png)

$A$: 10 events
5 bins (features)

$U$, $\Sigma$, $V$
Principal component analysis -- math

Now, the $i$th event can be decomposed as summation of eigenvectors $z_j$:

$$A_{i,:} \approx \sum_{j=1,2,\ldots,k} \sigma_j u_{i,j} v_{:,j}$$

$k$ is the cut we choose to drop out minor modes.

Because we take advantage of collective modes in all events!
PCA in Heavy-Ion

- subleading modes of factorization breaking

- Best linear descriptor

\[ \zeta_{n,\text{pred}}^{(a)} = \varepsilon_{n,n} + c_1 \varepsilon_{n,n+2} \]

- Nonlinear response coefficients

Rajeev S. Bhalerao, Jean-Yves Ollitrault, Subrata Pal, Derek Teaney

- Experimental data

CMS collaboration, Phys.Rev. C96 (2017) no.6, 064902

Can a machine automatically discover flow?

arXiv: 1903.09833
Ziming Liu, Wenbin Zhao, **Huichao Song**
Principal Component Analysis (PCA) of Collective Modes in Relativistic Heavy-Ion Collisions
Q: What are good bases to decompose particle distribution?

Fourier Transformation?

What makes a good flow observable?

You are right. But, approximately.

Previous work utilizes Fourier Transformation in the $\phi$ direction:

$$\frac{dN}{dp} = \sum_{n=-\infty}^{+\infty} V_n(p) e^{in\phi} \quad p = (p_t, \eta)$$

PCA decomposes $V_n(p)$ into eigenmodes:

$$V_n(p) = \sum_{\alpha=1}^{k} \xi^{(\alpha)} V_{n}^{(\alpha)}(p)$$

However, we apply PCA directly to $dN/d\phi$ data without FT:

$$\frac{dN}{d\phi} = \sum_{\alpha=1}^{k} \xi^{(\alpha)} \left( \frac{dN}{d\phi} \right)^{(\alpha)}$$
Pb+Pb collisions at 2.76 A TeV

Trento initial model ➔ Vishnew Hydrodynamics ➔ iss particle sampling

No hadron rescattering or resonance decays to simplify problem settings.
PCA for flow analysis

Data sets: \( \frac{dN}{d\phi} \)

With PCA, each flow distribution is decomposed into superposition of eigenmodes.

\[
\frac{dN}{d\phi} = \mu + x_1 z_1 + x_2 z_2 + x_3 z_3 + \ldots
\]
Singular values

Degeneracy $\rightarrow$ Rotational symmetry

Eigenvectors $z$: correlations between features
Singular values $\sigma$: importance of eigenvectors
Eigenvectors/ Principal components

Elliptic flow  Triangular flow  ...... 

Machines can automatically discover flow without any guidance from human beings!
Define flow harmonics with PCA

\[ \frac{dN}{d\phi} = \mu + \sum_{i=1}^{k} x_k z_k \]

Event average comparisons

<table>
<thead>
<tr>
<th>n</th>
<th>( v_n' \times 10^2 )</th>
<th>( \bar{v}_n \times 10^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>6.03</td>
<td>6.08</td>
</tr>
<tr>
<td>3</td>
<td>2.57</td>
<td>2.53</td>
</tr>
<tr>
<td>4</td>
<td>1.21</td>
<td>1.25</td>
</tr>
<tr>
<td>5</td>
<td>0.57</td>
<td>0.66</td>
</tr>
<tr>
<td>6</td>
<td>0.26</td>
<td>0.37</td>
</tr>
</tbody>
</table>

Define flow harmonics with Fourier (Tradition)

PCA flow harmonics \( \approx \) Traditional flow harmonics

Event-by-event comparisons
Symmetric cumulants

Fourier: $SC_v(m, n) = \langle v_m^2 v_n^2 \rangle - \langle v_m^2 \rangle \langle v_n^2 \rangle$

PCA: $SC'_v(m, n) = \langle v'_m^2 v'_n^2 \rangle - \langle v'_m^2 \rangle \langle v'_n^2 \rangle$

Correlation between different harmonics decrease for PCA!
Pearson correlation between initial and final PCA and Fourier.
Pearson correlation between initial and final

20%-30% centrality data

Fourier:

PCA:

PCA has a more diagonal pattern!
Conclusion for paper 1

• Without defining Fourier bases, PCA can automatically discover flow.

• We use PCA bases to re-define flow harmonics and find that
  • PCA bases have less correlations between mixed flow harmonics.
  • Flow observables have a more diagonal pattern for PCA
The limitation of studying sub-leading flow with PCA

Recent work, paper in preparation
Ziming Liu, Arabinda Behera, Huichao Song, and Jiangyong Jia
Reconsideration on studying sub-leading flow with PCA
Principal component analysis of event-by-event fluctuations

PRL 114, 152301 (2015)
Rajeev S. Bhalerao, Jean-Yves Ollitrault, Subrata Pal, and Derek Teaney

**Single particle distribution**

\[
\frac{dN}{dp} = \sum V_n(p)e^{-in\phi}
\]

**Two-particle correlation**

\[
\langle V_{n\Delta}(p_a, p_b) \rangle = \langle V_n(p_a)V_n^*(p_b) \rangle
\]

\[
\downarrow \text{diagonalization}
\]

**Decompose** \( V_n(p_T) \) **with PCA modes**

\[
V_n(p_T) = \zeta_n^{(1)}V_n^{(1)}(p_T) + \zeta_n^{(2)}V_n^{(2)}(p_T) + \zeta_n^{(3)}V_n^{(3)}(p_T) + \cdots + \zeta_n^{(k)}V_n^{(k)}(p_T)
\]

\[
\downarrow \text{Leading flow}
\]

\[
\downarrow \text{Sub-leading flow}
\]

\[
\downarrow \text{Sub-sub-leading flow}
\]

\[
\downarrow \text{k-th mode}
\]
Principal component analysis of event-by-event fluctuations

PRL 114, 152301 (2015)

Rajeev S. Bhalerao, Jean-Yves Ollitrault, Subrata Pal, and Derek Teaney

Non-zero sub-leading flow?
Cause for Sub-leading flow

Aleksas Mazeliauskas, Derek Teany

$V_3(p_T)$
Radial excitation leads to sub-leading flow!

Sub-leading flow is theoretically important!!!
Experimental results

Phys. Rev. C 96, 064902
The CMS Collaboration

Are these results solid?
Think more about it.

Model setup:
- AMPT model
- 1M UCC events
- subevent method

(paper in preparation)
Ziming Liu, Arabinda Behera, Huichao Song, and Jiangyong Jia
Reproduce CMS results

Same $p_T$ cut and proper $\eta$ gap $|\Delta\eta| < 0.8$. Our model can properly reproduce the results @CMS.
How to choose $\rho_T$ bins?

Orthogonal on this large interval

Not orthogonal on this small interval

Drop one $\rho_T$ bin each time, and redo the PCA

PCA will try to mix these two modes!

Will the modes still be stable?
How to choose $p_T$ bins?

Different choice of $p_T$ bins can introduce systematic errors! PCA modes are sensitive to our choice of $p_T$ bins!
Normalization

Traditional PCA
\[ \langle V_n \Delta (p_a, p_b) \rangle = \langle V_n(p_a)V_n^*(p_b) \rangle \]

New PCA
\[ \langle V_n \Delta (p_a, p_b) \rangle = \langle V_n(p_a)V_n^*(p_b) \rangle \]

\[ V_n(p_T) = \zeta_n^{(1)} V_n^{(1)}(p_T) + \zeta_n^{(2)} V_n^{(2)}(p_T) + \zeta_n^{(3)} V_n^{(3)}(p_T) + \cdots + \zeta_n^{(k)} V_n^{(k)}(p_T) \]

\[ \nu_n^{(\alpha)}(p_T) = \frac{V_n^{(\alpha)}(p_T)}{\langle M(p_T) \rangle} \]

\[ \tilde{\nu}_n(p_T) = \frac{\langle V_n \Delta (p_a, p_b) \rangle}{\langle M(p_a) \rangle \langle M(p_b) \rangle} \]

\[ \tilde{\nu}_n(p_T) = \frac{V_n(p_T)}{\langle M(p_T) \rangle} \]

\[ \tilde{\nu}_n(p_T) = \zeta_n^{(1)} \tilde{\nu}_n^{(1)}(p_T) + \zeta_n^{(2)} \tilde{\nu}_n^{(2)}(p_T) + \zeta_n^{(3)} \tilde{\nu}_n^{(3)}(p_T) + \cdots + \zeta_n^{(k)} \tilde{\nu}_n^{(k)}(p_T) \]

Different?
But which scheme can reveal more physics about initial profiles?
More analysis should be done to fully unravel the mystery of PCA!
Conclusion for paper 2

• The choice of $p_T$ bins introduces systematic errors, but we have no guidance from physics about how to choose them.

• Technically, the normalization procedure before/after PCA also lead to different results. Which is the real physics? Need more discussion.
Summary & Outlook
Principal Component Analysis:
- Unsupervised learning (dimensionality reduction)
- Good at discovering hidden correlations in data

PCA for flow discovery
- Integrate $p_T$, so stable
- Discover flow automatically
- Reduce mode coupling

PCA for sub-leading flow
- $p_T$ differential, sensitive to $p_T$ bins
- Ambiguity in normalization

All in all, PCA is a transparent yet powerful machine learning tool to extract main information/hidden correlations in data! But we should also be more careful about its results.

Outlook
Can PCA detect modes or structures from the massive data that is not realized or easily defined by human being?
Thank you for your attention!
VI. Back up
Model Details

- 2.76 A TeV Pb+Pb
- Viscous Hydro: VISH2+1
- EOS: s95-PCE
- Initial condition: TRENTo
- Hydrodynamic starting time $\tau_0 = 0.6 fm/c$
- Decoupling temperature $T_{sw} = \frac{148 MeV}{c}$
- $0.3 < p_T < 3 GeV$, Pions only
- 0%-10%,......,50%-60% totally 6 centrality bins, 2000 events for each bin
PCA Implementation

• Python – sklearn
• From sklearn.decomposition import PCA
• Mode cut \( k = 12 \)
Signal & Noise

Copper-Fryer:

\[
\frac{dN}{dy p_T dp_T d\phi} = \int_\Sigma \frac{g}{(2\pi)^3} p^\mu d^3 \sigma_{\mu} f(x, p)
\]
Bases Mixing (2\textsuperscript{nd} and 4\textsuperscript{th})

- Left: Fourier, $v_4 \sim \varepsilon_2$
- Right: PCA, $v'_4 \sim \varepsilon_2$
Bases Mixing (2\textsuperscript{nd} and 4\textsuperscript{th})

- Left: Fourier, \( \nu_4 \sim \varepsilon_4 \)
- Right: PCA, \( \nu'_4 \sim \varepsilon_4 \)
PCA for Initial Profiles

Smoothing Procedure

\[
\left( \frac{dS}{d\varphi} \right)_{\text{smooth}} = \int K(\varphi', \varphi) \frac{dS}{d\varphi'} d\varphi'
\]

\[
K(\varphi', \varphi) = \frac{1}{\sqrt{2\pi a}} e^{-\frac{(\varphi'-\varphi)^2}{2a^2}}, \quad a = 0.251 \text{ rad}
\]

For initial states, PCA=Fourier
Angle shift

Event plane Angle shift (alignment)

\[ \sigma_j \]

- No shift (Reaction Plane)
- Angle Shift (Align)
- Random Shift
Angle Shift

- The PCA is implemented in the reaction plane, so that eigenvectors mix 2\textsuperscript{nd} and 4\textsuperscript{th} flow harmonics.
- If randomly shifting every event (as in experiments), the bases will be exact Fourier bases due to rotational symmetry.
Good about Neural Networks

**Universal Approximation Theorem:**
A neural network with one layer but infinitely many neurons can approximate any continuous function.

\[ x \rightarrow f(x) \]

**Width v.s. Deep Theorem:**
An exponentially number of neurons in one layer = linear number of layers

\[ 2^d \text{ neurons} \quad \Rightarrow \quad d \text{ layers, } 2d \text{ neurons} \]

That is why we need DEEP learning!
How to train a model?
Train/Evaluate a model

- The evaluation is to evaluate the difference between the network’s outputs and learning targets. — loss function

For supervised learning, since there has been a target $y(x)$, loss function can be defined as

- $\ell(\theta) = \frac{1}{2n} \sum_x [y(x) - \hat{y}(x)]^2$
- $\ell(\theta) = -\frac{1}{n} \sum_x [y(x) \ln \hat{y}(x) - (1 - y(x)) \ln(1 - \hat{y}(x))]$

Train a model $\Leftrightarrow$ Minimize the loss function

Stochastic Gradient Descent

$$\theta' = \theta - \epsilon \frac{\partial \ell(\theta)}{\partial \theta}$$

What is Machine Learning?

**Machine learning (ML)** is the scientific study of algorithms and statistical models that computer systems use to effectively perform a specific task without using explicit instructions. ——[wikipedia](https://en.wikipedia.org/wiki/Machine_learning)

Linear Regression is also Machine Learning!
How PCA bases mix Fourier bases?
Bases Mixing (2\textsuperscript{nd} and 4\textsuperscript{th})

\[
\begin{pmatrix}
  z_1 \\
  z_2 \\
  z_5 \\
  z_6
\end{pmatrix} = 
\begin{pmatrix}
  a_{11} & a_{12} & a_{13} & a_{14} \\
  a_{21} & a_{22} & a_{23} & a_{24} \\
  a_{31} & a_{32} & a_{33} & a_{34} \\
  a_{41} & a_{42} & a_{43} & a_{44}
\end{pmatrix} 
\begin{pmatrix}
  \cos(2\phi) \\
  \sin(2\phi) \\
  \cos(4\phi) \\
  \sin(4\phi)
\end{pmatrix}
\]

If the eigenmodes of PCA is the same as fourier bases, the mixing matrix \( A \) should be identity. But actually, the matrix is not diagonal. Take data from centrality 30\% – 40\% for example

\[
A = 
\begin{pmatrix}
  0.956 & 0.295 & 0.213 & 0.053 \\
  -0.295 & 0.956 & -0.057 & 0.215 \\
  -0.217 & -0.041 & 0.960 & 0.209 \\
  0.035 & -0.215 & -0.219 & 0.951
\end{pmatrix}
\]
Bases Mixing (2\textsuperscript{nd} and 4\textsuperscript{th})

It is interesting to find that the mixing matrix $A$ follows the form below for all centrality classes. The parameters do not hold, but the form does.

$$A = \begin{pmatrix}
\cos(\theta_2) & \sin(\theta_2) & a\cos(\theta_2) & a\sin(\theta_2) \\
-sin(\theta_2) & \cos(\theta_2) & -a\sin(\theta_2) & a\cos(\theta_2) \\
-a\cos(\theta_4) & -a\sin(\theta_4) & \cos(\theta_4) & \sin(\theta_4) \\
asin(\theta_4) & -a\cos(\theta_4) & -\sin(\theta_4) & \cos(\theta_4)
\end{pmatrix}$$

To make notations easier, we denote

$$U_1 = \begin{pmatrix}
\cos(\theta_2) & \sin(\theta_2) \\
-sin(\theta_2) & \cos(\theta_2)
\end{pmatrix}, \quad U_2 = \begin{pmatrix}
\cos(\theta_4) & \sin(\theta_4) \\
-sin(\theta_4) & \cos(\theta_4)
\end{pmatrix}$$

Note that $U_1$ and $U_2$ are just rotation matrix in 2-d Cartesian coordinate.
Bases Mixing (2\textsuperscript{nd} and 4\textsuperscript{th})

\[ A = \begin{pmatrix} U_1 & aU_1 \\ -aU_2 & U_2 \end{pmatrix} \]

It is interesting to note that \( A \) can be decomposed into multiplication of simpler matrices.

\[ A = \begin{pmatrix} U_1 & 0 \\ 0 & I_2 \end{pmatrix} \begin{pmatrix} I_1 & 0 \\ 0 & U_2 \end{pmatrix} \begin{pmatrix} I_2 & a \\ -a & I_2 \end{pmatrix} \]

If we see the matrix as an operation, then the operation was decomposed into three steps:

- First, PCA mixed 2nd harmonic flow and 4th harmonic flow, by adjusting \( a \).
- Second, PCA mixed within 4\textsuperscript{th} order plane by adjusting \( \theta_4 \).
- Third, PCA mixed within 2\textsuperscript{nd} order plane by adjusting \( \theta_2 \).