Modified independent component analysis for extracting Eigen-Modes of a quantum system

Yadong Wu and Hui Zhai
Institute for Advanced Study, Tsinghua University, Beijing, 100084, People’s Republic of China

Keywords: independent component analysis, quantum dynamics, Bose–Einstein condensate

Abstract
Dynamical probes are popular experimental tools for studying quantum systems nowadays. Though a well-defined eigen-mode will manifest itself as a single frequency oscillation in the dynamical probe, the real experimental data can be quite messy because a probe usually excites multiple modes which are also mixed together with random white noise. Moreover, the oscillations usually quickly damp out due to finite decoherence time. These make it difficult to extract the frequencies of these oscillation measurement data, that is, the eigen-energies of these eigen-modes. Here we develop an unsupervised machine learning algorithm to solve this problem. Our method is inspired by the independent component analysis method and its application to the ‘cocktail party problem’, where the goal is to recover each voice from detectors that detect signals of many mixed voices. We demonstrate the advantage of our method by an example of analyzing the collective mode of a Bose–Einstein condensate of atomic gases. We believe that this method can find broad applications in analyzing data of dynamical experiments in quantum systems of different fields.

Introduction

In quantum systems, eigen-modes have discrete eigen-energies, and each eigen-mode with a fixed eigen-energy manifests itself as a single-frequency oscillation in a dynamical probe. The frequency of the oscillation, up to an \( h \), is the eigen-energy of the eigen-mode. Therefore it is quite a common task in many quantum physics experiments to extract a single-frequency oscillation from dynamical measurements.

If the measurement can be performed for a sufficiently long time, the frequency can be determined by the Fourier transformation. However, in many cases it is not possible to perform a long time measurement, either because the oscillation quickly damps out because of decoherence, or because the lifetime of the quantum system itself is limited. Hence, the frequency resolution of the Fourier transform method is always limited by \( 2\pi/T \), where \( T \) is the total duration for taking data. An alternative way commonly used in practice is to fit the dynamics with a damped harmonic oscillator. Nevertheless, it is quite often that the probe excites multiple modes with different frequencies, and some times the oscillations are also embedded in noisy signals, therefore the fitting becomes not so reliable. Thus, when the measurement time is short, and the data contains multiple frequencies and is noisy, the task of extracting single-frequency oscillation becomes quite challenging. The purpose of this work is to design a machine learning algorithm to solve this problem.

To start with, we point out that this problem bears a lot of similarities with the famous ‘cocktail party problem’ that has been widely studied in neuroscience and computer science for decades, and the independent component analysis (ICA) is a quite efficient algorithm to solve the cocktail party problem. Let us first briefly review the ICA method [1] and its application to the classical ‘cocktail party problem’ [2].

The ‘cocktail party problem’ considers \( N \)-number of sources of voices, each of which is a sequence of data denoted by \( s_i(t), (i = 1, \ldots, N) \) and \( M \)-number of detectors with \( M > N \), each of which obtains a signal \( x_j(t), (j = 1, \ldots, M) \) as a linear combination of all \( s_i(t) \). That is to say, there is a \( M \times N \) matrix \( A \) such that

\[
x_j(t) = \sum_i A_{ji} s_i(t).
\]

© 2020 The Author(s). Published by IOP Publishing Ltd
The matrix $A$ is unknown and is assumed to be independent of $t$, and the goal is to find out $A^{-1}$ such that one can recover each voice $s_i(t)$ from the signals obtained by the detectors $\{x_j(t), (j = 1, \ldots, M)\}$. The aforementioned problem in a quantum system is actually the same as the 'cocktail party problem'. Each single frequency oscillation from a given eigen-mode can be viewed as a source $s_i(t)$, and several different measurements $x_j(t)$ obtain several different superpositions of $s_i(t)$. To extract the frequency or eigen-energy of each mode, one first needs to recover $s_i(t)$ from the measurements $\{x_j(t)\}$.

What lies behind the ICA application of this problem is the central limit theorem $[3]$. For each signal, when we perform statistics over a certain duration of $t$, we can obtain the distribution for each signal, which always possesses certain features. The central limit theorem says that when one adds up many such signals together, the distribution of $x_j(t)$ will approach a Gaussian distribution and looks like a noise. Thus, if the inverting the transformation does not yield a single independent signal, in other words, if the inverting transformation still yields a summation of different signals, the distribution is still a Gaussian distribution. Therefore, in order for the inverted signal to be an independent signal, the distribution should be far away from Gaussian distribution as possible. The signal with Gaussian distribution displays maximum entropy for given mean and variance. Thus, the working principle of ICA is to find out $A^{-1}$ such that the distribution of $s_i(t) = (A^{-1})_i x_j(t)$ deviates from a Gaussian distribution as far as possible, or in a more quantitative description, that the entropy of $s_i(t) = \sum_j (A^{-1})_{ij} x_j(t)$ is as small as possible.

To be more precise, we first normalize the data $x_j(t)$ such that its mean value $\bar{x}_j(t) = 0$ and its covariance matrix $X'X = I$, where the statistics is performed over a sufficiently long duration of $t$. We further require $A' A = I$, which ensures the mean value of $s_i(t)$ is also zero and its variance also equals one. By performing statistics over $t$, for a given sequence $s(t)$, one can obtain its distribution $P[s, u]$, where $u$ is the range of $s(t)$. Then, we can compute the entropy of a given sequence $s(t)$ as

$$\mathcal{H}[s] = - \int P[s, u] \log P[s, u] du.$$  (2)

If $s_{\text{Gau}}$ is a sequence obeying a Gaussian distribution with zero mean and unity variance, $H[s_{\text{Gau}}] = \frac{1}{2} - \frac{1}{2} \log \frac{1}{2\pi}$ is the entropy maximum for sequences with same mean and variance. Hence, the ICA method is to find out $A^{-1}$ such that for each $s_i(t) = \sum_j (A^{-1})_{ij} x_j(t)$,

$$f[s] = \mathcal{H}[s_{\text{Gau}}] - \mathcal{H}[s]$$  (3)

is maximized. In practices, since it is hard to directly compute $P[s, u]$, several formulae have been proposed to approximate $\mathcal{H}[s][s]$.

**Results**

When we apply the ICA method to analyze the simulated experimental results of a dynamical probe, we will see that the ICA method can work but the outcome is not ideal. The reason that it does not work well can also be understood. Because the ICA only assumes that the signal has a certain feature but does not fully explore what exactly the feature is. This is good for the original 'cocktail party problem' because it does not require prior knowledge of each voice. However, as discussed above, in our quantum problem each eigen-mode has a fixed energy and manifests as a single frequency oscillation, but this feature of being single frequency oscillation is not utilized in the ICA method above. Therefore, the main result of this work is to present a modified ICA method that aims at finding out a signal of single frequency oscillation, short-noted as $s$-ICA.

**The loss function of $s$-ICA**

In short, let us consider a reference signal $s_{\text{ref}} = \sqrt{2} \cos(\omega t)$. Here $\sqrt{2}$ is chosen to ensure that the variance equals one. The ICA method is to find out an $A^{-1}$ such that each $s_i = \sum_j (A^{-1})_{ij} x_j$ is away from $s_{\text{Gau}}$ as much as possible; and our $s$-ICA method is to find out an $A^{-1}$ such that each $s_i = \sum_j (A^{-1})_{ij} x_j$ is close to $s_{\text{ref}}$ as much as possible.

To quantify how $s_i$ is close to $s_{\text{ref}}$ instead of using entropy we consider a quantity called the cumulants$[3]$. The cumulants is defined as

$$K[s, z] = \log(\langle e^{z s(t)} \rangle),$$  (4)

where $\langle \ldots \rangle$ means performing average over a certain duration of $t$. For $s_{\text{ref}}$, it is straightforward to compute

$$P[s_{\text{ref}}, u] = \frac{1}{\pi} \frac{1}{\sqrt{2 - u^2}},$$  (5)
for $-\sqrt{2} < u < \sqrt{2}$, and otherwise $\mathcal{P}[s_{\text{ref}}, u] = 0$, and then

$$
\langle e^{2 \pi i t u} \rangle = \int \mathcal{P}[s_{\text{ref}}, u] e^{2 \pi i u} du = I_0(\sqrt{2} z),
$$

where $I_0$ is the Bessel function of the first kind, and

$$
K'[s_{\text{ref}}, u] = \log(I_0(\sqrt{2} z)).
$$

It is very important to acknowledge that $\mathcal{P}[s_{\text{ref}}, u]$, and consequently, $K'[s_{\text{ref}}, u]$, is independent of the frequency $\omega$. Thus, we can use $K'[s_{\text{ref}}, u]$ as a reference to quantify how close $s_i$ is to $s_{\text{ref}}$ without knowing the value of $\omega$ as a prior. In principle, we should require $K[s, z]$ to be close to $K'[s_{\text{ref}}, z]$ for all $z$. In practices, we consider a set of $\{z_i\}$, $i = 1, \ldots, L$, and define the loss function as

$$
\mathcal{L}[s] = \sum_{i=1}^L (K[s, z_i] - K'[s_{\text{ref}}, z_i])^2.
$$

Therefore, our s-ICA method is to find out $A^{-1}$ such that $\mathcal{L}[\sum_i (A^{-1})_i x_i]$ is minimized.

**Iteration in s-ICA**

There is one subtlety in s-ICA. The analytical form equation (7) crucially relies on the distribution function equation (5) for $s_{\text{ref}}$, but equation (5) is correct only when the statistics is performed over a duration that is an integer times of the period $2\pi/\omega$. Thus, when we calculate $K[\sum_i (A^{-1})_i x_i, z]$, the statistics for all $x_i$ also needs to be carried out for the time interval being integer times of $2\pi/\omega$, otherwise $K[s, z]$ always can not perfectly converge to $K'[s_{\text{ref}}, z]$. Hence, it enters a paradox. Since the goal is to separate out a single frequency oscillation to determine the frequency $\omega$, $\omega$ is not known before analyzing. To solve this problem, our s-ICA method requires performing the minimization of $\mathcal{L}[s]$ iteratively. That is to say, we first choose an arbitrary period $r$ to perform statistics for $x_i$, with which we minimize $\mathcal{L}[\sum_i (A^{-1})_i x_i]$ to find out a $s^{(1)}_i(t)$. Although $s^{(1)}_i(t)$ is not a perfect single frequency oscillation because of this reason, we can still roughly determine a frequency $\omega^{(1)}_i$ by fitting the results with a single frequency sinusoidal oscillation. Then we perform statistics for $x_i$ with time duration $2\pi/\omega^{(1)}_i$ and minimize $\mathcal{L}[\sum_i (A^{-1})_i x_i]$ to find out a $s^{(2)}_i(t)$. Because this time, the statistics is performed in a time interval more close to integer times of the actual period, the cumulants is more close to the reference function equation (7). Therefore, $s^{(2)}_i(t)$ will be more close to a single frequency oscillation than $s^{(1)}_i(t)$, from which we can determine a frequency $\omega^{(2)}_i$. We can continue the procedure and the accuracy of the frequency will be continuously improved, until we obtain a very good single frequency signal after $k$-steps and the frequency determined at each step also converges. This completes our s-ICA method. A comparison between ICA and our modified s-ICA method is shown in figure 1.

**Analyzing collective modes of a BEC**

As a demonstration of our method, we apply our method to analyze the collective oscillations of a Bose–Einstein condensate (BEC) of cold atoms in a harmonic trap, which is a very common experimental measurement in cold atom experiments[6–9]. We use this as an example simply because it is easy to generate simulated experimental data in this case. We first consider an equilibrium Thomas-Fermi density distribution

$$
n_0(r) = \frac{\mu - V(r)}{g} \Theta(n_0(r)),
$$

Table 1: A comparison between ICA method and our modified ICA method, in term of the goal, the loss function and the time period for performing statistics of the data.

|          | ICA                  | s-ICA                |
|----------|----------------------|----------------------|
| Goal     | Extracting source with certain feature | Extracting source that is a single frequency oscillation |
| Loss Function | Minimizing entropy | Minimizing the cumulants with respect to the cumulants of single frequency oscillation |
| Statistics | Over sufficiently long time | Over integer times of period (reach by iteration) |
Figure 1. (a)–(c): The original data of density oscillation of a Bose–Einstein condensate from three different detectors. Three different eigen-modes obtained by the ICA method (d)–(f), by our s-ICA method at the first round (g)–(i); and by the s-ICA method at the second round (j)–(l).

where $\Theta(x) = \max(0, x/|x|)$ is the Heaviside function, and for simplicity, we consider a two-dimensional geometry with a harmonic trap $V(r) = m\omega_{\perp}^2 (x^2 + y^2)/2$ and $\omega_{\perp}$ being the trapping frequency. $\mu$ is the chemical potential that later will be adjusted to satisfy the total number of atoms conservation condition at any given time, and the total number of atoms is chosen as $N = 10^3$ in our simulation. $g$ is the interaction strength, and we here set $\hbar = 1$, $m = 1$, $\omega_{\perp} = 1$ and $g = 10$. We consider that three different modes are excited, and they are

\begin{align}
\delta n_1 &= x \cos(\omega_1 t), \\
\delta n_2 &= (x^2 - y^2) \cos(\omega_2 t) \\
\delta n_3 &= (x^2 + y^2) \cos(\omega_3 t),
\end{align}

which are the dipole mode, the quadrupole mode and the breathing mode, with their frequencies being $\omega_1 = 1$, $\omega_2 = \sqrt{2}$ and $\omega_3 = 2$, respectively [10, 11]. We also add noise that varies at different spatial locations and different time, denoted by $\epsilon(r, t)$ and rate $\tau$ to mimic the damping situation in the experiments. Then the simulated data of the real time density dynamics is given by

\[
n(r, t) = \left( n_0(r) + e^{-t/\tau} \sum_{i=1}^{3} f_i \delta n_i + \epsilon(r, t) \right) \Theta(n(r, t)).
\]

where $f_i$ ($i = 1, 2, 3$) denotes the amplitudes of the three modes.

Here we shall also note that, although $x$, $x^2 - y^2$ and $x^2 + y^2$ initially are three functions orthogonal to each other, due to the boundary condition imposed by the Heaviside function and the atom number conservation condition, they are no longer orthogonal. Thus, the method like the principle component analysis [12] also does not work very well in this case (See supplementary material for details (stacks.iop.org/MLST/1/025010/mmedia)).

Now we will use the ICA method and our s-ICA method to separate out each $\delta n_i$ from data $n(r, t)$ given by equation (13). As an example, we choose $f_1 = 1$, $f_2 = 0.2$, $f_3 = 0.2$ and $\tau = 20$, and $\epsilon$ uniformly distributed in the range $[-2, 2]$. Here we only require short-time information of couple oscillation periods, and in fact, for this application another advantage is that we only need information from a few spatial points. Let us consider three locations denoted by $r_1$, $r_2$ and $r_3$, and the density dynamics at these points play the role as detectors. Their density dynamics all contain these three frequencies but the coefficients are different because it depends on the spatial locations. We plot $n(r_i, t)$ ($i = 1, 2, 3$) in figure 1(a)-(c), which show irregular temporal behaviors. The goal of ICA or s-ICA is to find out a proper combination of them that exhibits the single frequency oscillatory behavior.
In figure 1(d)-(f), we show the results of the ICA method. We can see that the quality of the results is not very good. In figure 2, we also show the cumulants calculated for the signal of figure 1(d), and one can clearly see that it derivates from $K_{\text{ref}}$. This means that although the signal obtained by the ICA possesses a certain feature, it is not yet a perfect single frequency oscillation.

In figure 1(g)-(i), we show the results of the s-ICA method. The results are also not perfect, and the cumulants for the signal of figure 1(g) is also shown in figure 2 to compare with $K_{\text{ref}}$, and one can see the discrepancy still exists. Here we use this fitting equation to determine the oscillation frequencies of analyzing result:

$$s_{\text{ICA}} = A_1 \cos(\omega t + \varphi) + A_2$$  (14)

where $A_1, A_2, \varphi, \omega$ are all fitting parameters and $\omega$ is what we want. Nevertheless, by fitting these signals, one can obtain $\omega_1^{(1)} = 1.068, \omega_2^{(1)} = 1.432$ and $\omega_3^{(1)} = 2.011$. Next, we perform statistics for two periods of time duration $4\pi/\omega_1^{(1)}$ and obtain the result shown in figure 1(j), with which we obtain a new frequency $\omega_1^{(2)} = 1.005$, and it only differs from the real value by 0.5%. This is significantly accurate than the limit of frequency resolution given by the Fourier transformation [supple]. For performing statistics, we have taken 400 points in this time interval. Similarly, we perform statistics for time duration of $4\pi/\omega_1^{(2)}$, or for time duration of $4\pi/\omega_2^{(1)}$, and obtain results shown in figure 1(k) and (l), respectively, with which we determine $\omega_2^{(2)} = 1.421$ and $\omega_3^{(2)} = 1.990$. These results are also very close to the actual value with only 0.48%, 0.5% deviation. The cumulants for the signal of figure 1(j) also agrees perfectly with $K_{\text{ref}}$, as shown in figure 2. Here we use the optimization processes to minimize the loss function equation (8) with Newton's method.

Here we should note that so far we only obtain the frequency but can determine neither the amplitude nor the damping rate. In fact, both ICA and our s-ICA have a problem that the amplitude can not be uniquely determined. In fact, signals figures 1(j)-(l) are all distributed between $[-\sqrt{2}, \sqrt{2}]$ because the data $n(r, t)$ ($i = 1, 2, 3$) have been preprocessed to be zero mean and unity variance. For instance, to obtain the amplitude, we will perform another fitting. Here we consider density dynamics from all spatial points without pretreatment, that is

$$n(r, t) = C_0(r) + \sum_{i=1}^{3} C_i(r) \cos(\omega_i t).$$  (15)

Now all $\omega_i$ are already known from the s-ICA analysis. For each $r$ point, there are only four fitting parameters $C_i(r)$ ($i = 0, \ldots, 3$) to fit data points of a sequence of $t$. The results of the fitting are shown in figure 3 which reveal the spatial information of each mode. In this case, figure 3(a) is the background. Figure 3(b) shows positive amplitude in one side and negative amplitude in another side, which changes the BEC’s center-of-mass and is the dipole mode. Figure 3(c) has two nodal lines and concentrates at the surface, and it is the quadruple mode. Figure 3(d) shows negative amplitudes at the center and positive amplitudes in outside, which changes the BEC’s side and is the breathing mode [11]. With similar fitting, we can also determine the damping rate of each mode.
Outlook

In summary, in this work we have developed a modified ICA method to extract the eigen-energy of a quantum system from a dynamical probe. We remark that our method and the Fourier transformation respectively work well in two different regimes. The Fourier transformation requires taking data over a sufficiently long time duration, but it does not require many data within each period. If the total time duration is short, the frequency resolution obtained from the Fourier transformation will be quite limited. On the contrast, our method only needs data for a short time duration such as couple periods, but we require the accumulation of sufficient data during the time interval of a few oscillation periods that allows performing accurate statistics. Hence, our method and the Fourier transformation are very complementary, and therefore we believe our method can find wide application in future data analysis. Although the system considered here is a quantum one, the data is the expectation value of certain observable and is a class one. Hence, our method may even find its application beyond physics problems.

Another remark is that it is of great interest to consider the quantum data and the quantum analogy of the 'cocktail party problem', and to see whether the similar ICA method can work there [13].

Acknowledgment

We thank Jia-Ming Li, Ning Sun and Ce Wang for helpful discussion. This work is supported MOST under Grant No. 2016YFA0301600, NSFC Grant No. 11734010 and Beijing Outstanding Scholar Program.

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID iD

Yadong Wu  https://orcid.org/0000-0002-4135-8813

References

[1] Hyvärinen A and Oja E 2000 Neural Netw. 13 411
[2] Girolami M and Fyfe C 1997 Neural Netw. 10 1607
[3] Bárány I and Vu V 2007 Ann. Probab. 35 1593
[4] Hyvärinen A 1997 New Approximations of Differential Entropy for Independent Component Analysis and Projection Pursuit - A47
[5] Lukacs E 1970 Characteristic Functions (New York: Hafner Publishing Company)
[6] Stringari S 1996 Phys. Rev. Lett. 77 2360
[7] Jin D S, Ensher J R, Matthews M R, Wieman C E and Cornell E A 1996 Phys. Rev. Lett. 77 420
[8] Jin D S, Matthews M R, Ensher J R, Wieman C E and Cornell E A 1997 Phys. Rev. Lett. 78 704
[9] Stamper-Kurn D M, Miesner H J, Inouye S, Andrews M R and Ketterle W 1998 Phys. Rev. Lett. 81 500
[10] Ho T-L and Ma M 1999 J. Low Temp. Phys. 115 61
[11] Pethick C J and Smith H 2008 Bose–Einstein Condensation in Dilute Gases Section 7.3 (Cambridge: Cambridge University Press)
[12] Jolliffe I T 2002 Principal Component Analysis (New York: Springer)
[13] Liang X, Wu Y and Zhai H 2019 arXiv: 190406411