ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-MOPAB215

USING ICA FOR RETRIEVING TENG PARAMETERS

A. Lauterbach*, G. Franchetti[1, 2], Institute for Applied Physics
Goethe University, Frankfurt, Germany
[1]also at GSI Helmholtzzentrum für Schwerionenforschung GmbH, Darmstadt, Germany
[2]also at Helmholtz Research Academy Hesse for FAIR (HFHF)

Abstract

The blind source separation (BSS) method of Independent Component Analysis (ICA) is explored as a new approach for the reconstruction of the transfer matrix of Linear Coupling Parameterization. ICA is a method to detangle independent signals out of several measurements of their mixtures. In BSS-calculations, it is usually not possible to retrieve the mixing matrix, for the source signals, as well as the matrix, are unknown. Combining the parameterization model of D.A. Edwards and L.C. Teng with the standard ICA approach, it is though possible to retrieve the mixing matrix, as the form of the original uncoupled motion is known. At the same time arises the possibility to recalculate the parameters of Edwards and Teng through a system of equations of the one turn map components. It can be shown as a proof of concept, that the parameters can be reconstructed up to high accuracy for a simulated, non-perturbed signal.

LINEAR COUPLED MOTION

The evolution of a particle with coordinates (x, x', y, y') is determined by the transfer map of each accelerator element. After one turn the particle encounters the same sequence of elements, and in order to determine the dynamics, it is relevant to relate the optics parameters to the elements of the one turn map. If the elements are decoupled, the one turn map is the direct product of two 2×2 matrices, one for each plane. The presence of linear coupling elements, as skew quadrupolar errors, alter the usual 2×2 description and the parameterization of motion gets more complex. The one turn map becomes a 4×4 matrix, whose elements provide information on the feature of the particle motion.

PARAMETERIZATION

D.A. Edwards and L.C. Teng extended the parameterization of Courant and Snyder [1] to a two-dimensional system with a 4x4 matrix [2]. Using the condition of simplecticity and Teng's previous work about parameterization of symplectic matrices to discover independent quantities [3], they define the matrix for the transport of a particle from the longitudinal coordinate z_1 to z_2 as RUR^{-1} with

$$R = \begin{pmatrix} I\cos\phi & D^{-1}\sin\phi \\ -D\sin\phi & I\cos\phi \end{pmatrix}, \quad U = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \tag{1}$$

with

$$D = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{2}$$

$ad - bc = 1$ and with ϕ the coupling angle. Here I is the 2×2 unit matrix, A_1 and A_2 are the corresponding Courant-Snyder parameterizations for the respective coordinates z_1 and z_2 [1]:

$$A_{1/2} = I\cos\mu_{1/2} + \begin{pmatrix} \alpha_{1/2} & \beta_{1/2} \\ -\gamma_{1/2} & -\alpha_{1/2} \end{pmatrix} \sin\mu_{1/2}, \tag{3}$$

with μ_1, μ_2 being the phase advances. Furthermore, Edwards and Teng presented the approach of a canonical transformation from a coupled motion $\vec{X}$ to an uncoupled motion $\vec{V}$ using the inverse of the matrix R:

$$\vec{V} = R^{-1}\vec{X} \tag{4}$$

with $\vec{V}^T = (u, p_u, v, p_v)$ and $\vec{X}^T = (x, p_x, y, p_y)$. The derived parameterization of the vector $\vec{V}$ reads

$$\vec{V} = \begin{pmatrix} u \\ p_u \\ v \\ p_v \end{pmatrix} = \begin{pmatrix} \sqrt{W_1\beta_1}\cos(\Psi_1) \\ -\sqrt{W_1/\beta_1}(\sin\Psi_1 + \alpha_1\cos\Psi_1) \\ \sqrt{W_2\beta_2}\cos(\Psi_2) \\ -\sqrt{W_2/\beta_2}(\sin\Psi_2 + \alpha_2\cos\Psi_2) \end{pmatrix} \tag{5}$$

where Ψ_1 and Ψ_2 are the phases, $\beta_{1/2}, \alpha_{1/2}$ are the new optical functions and $W_{1/2}$ are the bilinear invariants

$$W_{1/2} = \vec{V}^T P^T \begin{pmatrix} J_{1/2} & 0 \\ 0 & 0 \end{pmatrix} \vec{V} \tag{6}$$

with P being the 4×4 unit symplectic matrix as definded in [2]. We next consider Eq. (4) with R corresponding to one turn, hence $\vec{X}$ being the particle coordinate at each turn.

LINEAR COUPLING AND ICA

In order to describe the coupled optics, Teng's parameters should be retrieved from the 4×4 one turn map. However, this task is difficult by measuring the coordinates of a beam freely oscillating in an accelerator. In the following we investigate the possibility of using ICA to ease this task. ICA is a BSS-method to reconstruct original data sets (signals) out of multiple measurements of mixed data sets (signals), implying that the original data sets are statistically independent of each other. The core problem of ICA can be formulated as $\vec{X} = M\vec{S}$, assuming the measured signals $\vec{X}$ to be weighted sums of the original signals $\vec{X}$, mixed by a mixing matrix M, such that $x_i = \Sigma_{j=1}^{N} m_{ij}s_j$. Each mixed data set is then a vector whose elements are weighted sums of the elements of the original data set, with m_{ij} being the weighting factors of the i components. To reconstruct the independent components s_i, the inverse of the mixing matrix M^{-1} needs to be found and

* annemarie.lauterbach@stud.uni-frankfurt.de

ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-MOPAB215

applied to the measurable data $\vec{X}$. The canonical transformation formalism of Edwards and Teng, displayed in Eq. (4), is similar to the basic ICA approach $\vec{X} = M\vec{S} \Rightarrow \vec{S} = M^{-1}\vec{X}$ with 1) $\vec{V}, \vec{S}$ being the original signals; 2) R, M being the mixing matrices; 3) $\vec{X}$ being the mixed, respectively coupled signals. Thus, we can treat the decoupled motion $\vec{V}$ as an original signal $\vec{S}$ that was mixed by a matrix M to form a signal $\vec{X}$.

PRINCIPLES OF ICA

Requirements to the Data Set and Data Preparation

A given data set, to be processed by the ICA algorithm, needs to meet the following requirements: 1) the original signals must be statistically independent, 2) only one of the original signals is allowed to be Gaussian and 3) the mixing matrix has to be square (which is the case having a 4×4 matrix). The condition of the source components $v_i(n)$ being independent from each other is equivalent with the condition of the autocovariance matrix $C = cov(\vec{v}_n \vec{v}_n) = E\{\vec{v}_n \vec{v}_n^T\}$ to be diagonal (here $E\{\cdot\}$ means average). The second condition is due to the fact, that within the algorithm chosen [4], non-Gaussianity is used as a measure of the independence of the rows of the demixing matrix with a random vector. Giving two Gaussian signals in the data set, the algorithm would only be able to reconstruct them as one signal. Before applying the actual ICA algorithm, the data needs to be prepared by centering and whitening. Centering is forcing the mean to be zero by $X_{i,\text{centered}} = X_i - \frac{1}{N}\sum_{n=1}^{N} X_i(n)$ and whitening is forcing the autocovariance matrix of the mixed signal $\vec{X}$ to be unity. Next by using a singular value decomposition of form $\vec{X}\vec{X}^T = UQZ^*$, we use the matrices Q and Z for creating a whitening matrix $M_{wh} = ZQ^{-\frac{1}{2}}Z^T$. With that we obtain a whitened mixed signal $\vec{X}_{wh} = M_{wh}\vec{X}$ that fulfills $E\{\vec{X_{wh}}\vec{X_{wh}}^T\} = I$.

Ambiguities

The method of ICA is intrinsically afflicted by limitations. The ambiguities of the outcome are: 1) Scaling: output values may be multiplied by a certain factor for each signal; 2) Permutation: the order of the signals in the output may differ from that in the input, and 3) Sign: output values may have opposite sign. Resolving these ambiguities is a necessary step to reconstruct the matrix elements of the one turn map and herewith the Teng parameters and is one of the challenges of the new concept introduced in this study.

PROOF OF CONCEPT

While the mixing/demixing matrix is and remains unknown in terms of the original ICA approach, combining it with the theory of Edwards and Teng and knowing the components v, p_v, u, p_u of $\vec{V}$ from [2], we can rewrite Eq. (4) as $\vec{V} = T\vec{S}$ with

$$T = \begin{pmatrix} \sqrt{W_1}F_1 & 0 \\ 0 & \sqrt{W_2}F_2 \end{pmatrix}, \qquad \vec{v} = \begin{pmatrix} \sin(\Psi_1) \\ \cos(\Psi_1) \\ \sin(\Psi_2) \\ \cos(\Psi_2) \end{pmatrix} \tag{7}$$

whereby

$$F_{1/2} = \begin{pmatrix} \sqrt{\beta_{1/2}} & 0 \\ -\alpha_{1/2}\frac{1}{\sqrt{\beta_{1/2}}} & -\frac{1}{\sqrt{\beta_{1/2}}} \end{pmatrix} \tag{8}$$

However, $\vec{V}$, in terms of R, $\vec{X}$, using Eqs. (1), (2) and the definition $\mathbf{c} = \cos$, $\mathbf{s} = \sin$, can also be formulated as

$$\begin{pmatrix} u \\ p_u \\ v \\ p_v \end{pmatrix} = \begin{pmatrix} \mathbf{c} & 0 & -d\mathbf{s} & b\mathbf{s} \\ 0 & \mathbf{c} & c\mathbf{s} & -a\mathbf{s} \\ a\mathbf{s} & b\mathbf{s} & \mathbf{c} & 0 \\ c\mathbf{s} & d\mathbf{s} & 0 & \mathbf{c} \end{pmatrix} \begin{pmatrix} x \\ p_x \\ y \\ p_y \end{pmatrix}. \tag{9}$$

The previous equation, when specialized to the one-turn map, reads

$$\begin{pmatrix} x \\ p_x \\ y \\ p_y \end{pmatrix}_n = \vec{X}_n = \begin{pmatrix} I\mathbf{c} & D^{-1}\mathbf{s} \\ -D\mathbf{s} & I\mathbf{c} \end{pmatrix} \begin{pmatrix} F_1 & 0 \\ 0 & F_2 \end{pmatrix} \vec{S}_n, \tag{10}$$

with $\vec{v_n}$ being dependent on the turn number n:

$$\vec{S}_n = \begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{pmatrix}_n = \begin{pmatrix} \sin(\Psi_1 + n\mu_2) \\ \cos(\Psi_1 + n\mu_1) \\ \sin(\Psi_2 + n\mu_2) \\ \cos(\Psi_2 + n\mu_1) \end{pmatrix}. \tag{11}$$

Here Ψ_1, Ψ_2 are phases determined by the initial conditions, and μ_1, μ_2 are the phase advances per turn. In the end, we have a problem mathematically formulated in a way that may be solvable with ICA, as long as the components s_i are independent from each other. Writing out the matrices I, D, E and F and using abbreviations: $\mathbf{s}_1 \equiv \sqrt{W_1}\sin\phi, \mathbf{s}_2 \equiv \sqrt{W_2}\sin\phi, \mathbf{c}_1 \equiv \sqrt{W_1}\cos\phi, \mathbf{c}_2 \equiv \sqrt{W_2}\cos\phi$ the product of the two matrices in Eq. (10) reads[1]

$$M \equiv RT = \begin{pmatrix} c_1\sqrt{\beta_1} & 0 & s_2\sqrt{\beta_2}\left(d + b\frac{\alpha_2}{\beta_2}\right) & s_2\frac{b}{\sqrt{\beta_2}} \\ -c_1\frac{\alpha_1}{\sqrt{\beta_1}} & -c_1\frac{1}{\sqrt{\beta_1}} & -s_2\sqrt{\beta_2}\left(c + a\frac{\alpha_2}{\beta_2}\right) & -s_2\frac{a}{\sqrt{\beta_2}} \\ -s_1\sqrt{\beta_1}\left(a - b\frac{\alpha_1}{\beta_1}\right) & s_1\frac{b}{\sqrt{\beta_1}} & c_2\sqrt{\beta_2} & 0 \\ -s_1\sqrt{\beta_1}\left(c - d\frac{\alpha_1}{\beta_1}\right) & s_1\frac{d}{\sqrt{\beta_1}} & -c_2\frac{\alpha_2}{\sqrt{\beta_2}} & -c_2\frac{1}{\sqrt{\beta_2}} \end{pmatrix} \tag{12}$$

Requirements to the Input Data for ICA

To apply ICA successfully, the source signals need to be statistically independent from each other. This is equivalent with the condition of the autocovariance matrix C with $C = cov(\vec{S}_n\vec{S}_n) = E\{\vec{S}_n\vec{S}_n^T\}$ to be diagonal. As ICA is a statistical tool, the results are the more reliable the larger the measured data set is.

[1] In the context of Eq. (12), s_1 and s_2 are temporarily used abbreviations for sinus terms to arrange the formula more clearly. These abbreviations are explicitly not the source signals.

ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-MOPAB215

RECONSTRUCTION OF TENG PARAMETERS

For our test we adopt the model of Eq. (10), in which the quantity $\vec{X}$ is our observable measured at turn n. In this model, the source signals s_i with $i = 1, 2, 3, 4$ are *"statistically independent"*, as long as $\mu_1 \neq \mu_2$ and for a large enough number of turns N aquired. We perform a test case in which we consider the independent signals $\vec{S}$ as in Eq. (11), which represent the uncoupled motion. The "mixed signals" $\vec{X}$, which result from measurements, are instead obtained from Eq. (10), respectively Eq. (12), in consistency with the general ICA approach of $\vec{X} = M\vec{S}$. The parameter values used for the test case's one turn map are displayed in Table 1 (column "orig"). The choice of the initial phases Ψ_1 and Ψ_2 in Eq. (11) is related to the initial condition of the test particle, which in this test do not play a significant role. From here, we will notate a quantity O as O' when reconstructed by the ICA algorithm. Applying the ICA algorithm on $\vec{X}_n$, we expect to be able to reconstruct $\vec{S'}_n$ and M' via

$$\vec{S'}_n = M'^{-1}\vec{X}_n. \tag{13}$$

To perform the test we consider $N = 3 \times 10^5$ turns and we choose a FASTICA algorithm as described in [4], which shows high performance while allowing economical hardware utilization. This algorithm maximizes the independency of s'_j (j = 1,2,3,4) to one another, and at the same time forms a Matrix M', so that Eq. (13) holds.

Dealing with the Ambiguities, Mapping

The ICA ambiguity of scaling can easily be solved, knowing that the maxima of the s'_j have a value of 1 (see Eq. (11)). We can therefore set rescaling factors $f_j = \frac{1}{|\max(s'_j)|}$.

The rescaling also needs to be applied to the corresponding rows of M'. To know, which are the corresponding rows we need to resolve the permutation ambiguity of ICA in the next step, identifying what is the permutation responsible of swapping the ordering of the components of $\vec{S'}$ in comparison to the components of $\vec{S}$. Due to intrinsic properties of ICA, it does not necessarily mean that s'_j corresponds to s_i when $i = j$. This means that a permutation-mapping of the original signals s_i to the corresponding reconstructed signals s'_j has to be found. We find this correspondence by using the method of cross-correlation based on [5] to evaluate the pair-wise similarity of each possible set of original signals s_i and reconstructed signals $\pm s'_j$. We pick the pairs with the best matches, including the sign, and so we identify the permutation matrix M_p so that $\vec{S'} = M_p\vec{S}$, which also resolves the ambiguity of the sign issue. At this point, M_p and the factors f_j can be applied to the corresponding rows of the reconstructed matrix M' to eliminate the ambiguities. We define a matrix M_{pf} as the component-wise product of the permutation matrix M_p of and the inverse of f_j: $(M_{pf})_{jm} \equiv \frac{1}{f_j}(M_p)_{jm}$, where the index m notates the columns of the matrix. We call $M'_{pf} = M_{pf}M'$ the rescaled and repermutated version of M'.

Final Reconstruction

With the permutation, sign and scaling problems solved, we can proceed, undoing the changes that the preparation step of whitening did to the turn-by-turn "measured" coordinates $\vec{X}_n$. Hence the reconstructed map $M_{rec} = M_{wh}{}^{-1} M'_{pf}$ is retrieved. Once M_{rec} is calculated, the parameters α_1, α_2, β_1, β_2, W_1, W_2, a, b, c and d can be retrieved by a system of equations with the help of Eq. (12), setting $M_{rec} \stackrel{!}{=} M$. The reconstruction results for the test example we discuss here are shown in Table 1, also displaying the relative error $\Delta_{rel} = [(\cdot)_{rec} - (\cdot)_{orig}]/(\cdot)_{orig}$.

Table 1: Reconstruction Results, Simulated for $N = 3 \times 10^5$ Turns And $\phi = 0.10258$, $\Psi_1 = 2.0$, $\Psi_2 = 0.1$, $\mu_1 = 0.36398$, $\mu_2 = 0.25314$

Parameter	**orig**	**rec**	Δ_{rel}
α_1	1.8739	1.8721	9.6173e-04
α_2	0.2636	0.2675	0.0148
β_1	16.8490	16.8363	7.5163e-04
β_2	13.6740	13.6870	9.4830e-04
W_1	0.5000	0.5000	3.5659e-05
W_2	0.9000	0.9000	1.1916e-05
a	1.4313	1.4317	2.5042e-04
b	4.5003	4.4644	0.0080
c	-0.0093	-0.0093	5.0225e-04
d	0.6694	0.6695	1.2004e-04

The results of Table 1 shows that the method is able to provide reconstructed parameters with a deviation from the original input values between 1.48% or less for the example discussed here.

CONCLUSION

As the Edwards-Teng-Parameterization leads to a natural formulation of the dynamics formally equivalent to that of an ICA algorithm, we have explored the possibility of using ICA for retrieving the Teng's parameters. The known structure of the parameterization allows to retrieve the Teng Parameters out of the reconstructed mixing matrix M with good accuracy for the example discussed. The reconstruction is, however, only possible when a successfull mapping of original and reconstructed signals is achieved solving the ICA intrinsinc ambiguities. We leave to future studies the investigation of the influence of the number of turns N needed and the performance of the algorithm with different values for μ_1 and μ_2.

REFERENCES

[1] E. D. Courant and H. S. Snyder, "Theory of the alternating-gradient synchrotron", *Annals of Physics*, vol. 281, pp. 360-408, Apr. 2000. doi:10.1006/aphy.2000.6012

ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-MOPAB215

[2] D. A. Edwards and L. C. Teng, "Parametrization of Linear Coupled Motion in Periodic Systems", *IEEE Transactions on Nuclear Science*, vol. 20, no. 3, pp. 885-888, Jun 1973. doi:10.1109/TNS.1973.4327279.

[3] L. C. Teng, "Concerning n-Dimensional Coupled Motions", FNAL, Batavia, Illinois, United States, Rep. FERMILAB-FN-0229, May 1971.

[4] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, NY, New York, John Wiley & Sons, 2001.

[5] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, Cambridge, MA, USA: MIT Press, 2016.