Harmonic Elimination of Inverters using Blind Signal Separation

B. Justus Rabi and R. Arumugam
College of Engineering, Anna University, Chennai-25, India

Abstract: Harmonics cause major problems in power system due to non linear loads. With the rapid growth in the utilization of the rectifier for critical loads, e.g. In computer or medical equipment the need for high quality uninterruptible power is increasing. The non-linear current drawn by the rectifier loads distort the output of UPS system. Mainly because of output filter impedance in UPS. A distorted output voltage will result in reduced dc-bus voltage or rectifier loads and may lead to excessive losses and heating. The main objective of this study is to eliminate harmonics and reduce the power loss in inverters. The harmonic separation process is implemented with a processor achieves low THD using Blind Signal separation. It is mostly used in medical instrumentation and medical applications like ECG, EEG.

Key words: Independent component analysis (ICA), blind signal separation (BSS), total harmonic distortion (THD), uninterrupted power supply (UPS)

INTRODUCTION

Power quality is of increasing concern to utilities and their customers alike[1]. In an optimum operation of power system, it is desirable to have voltage and current waveform as close to sinusoids as possible. Quality power supply means maintaining voltage and current as its rated rms with negligible amount of harmonics. There is a large growth in the use of loads that can be sensitive to power quality disturbances. The harmonic interferences in the wer systems, which are caused by non-linear loads, have been serious problems to solve. Blind Signal Separation is the best way to eliminate harmonics in inverters.

Most electronic equipment in the workplace raises likelihood of the potential interactions with the electric distribution system and requires a more sophisticated approach to preventing these interactions. Common power quality concerns, including voltage sags, swells and surges have led to the increased use additional facility equipment, such as uninterruptible sources and battery supported systems, to increase electrical reliability .In addition, signal interactions in sensitive equipment can be difficult to trace. Using the neural networks the mixed signals can be separated by repetitive training of neural network.

Neuro computing: Neuro computing[2] is concerned with processing information. The Neuro computing approach to information processing involves a learning process within an artificial neural network architecture that adaptively responds to inputs according to a learning rule. After the neural network has learned what it needs to know, the trained network can be used to perform certain tasks depending on the particular application. The prime reason for its popularity, when compared to other artificial intelligent paradigms are as follows:

The ability to recall a stored memory given only partial input data. The ability to learn by example and generalize. Fault tolerant, robust and distributed in nature. Non-linearity and Adaptivity. The ability to represent unclear, non-deterministic assignments, which cannot be described by the conventional logical programming languages[3].

In the future, artificial neural networks[4] are expected to solve the problem that could not be dealt with, using the conventional Von Neumann computers such as Feature extraction and blind signal separation.

Independent component analysis: A recent trend in neural network research is to study various forms of unsupervised learning beyond standard Principal Component Analysis (PCA) to solve nonlinear problems. Independent Component Analysis (ICA) is a useful extension of PCA. Independent Component Analysis is an unsupervised learning technique that in many cases characterizes the data in a natural way. The main application area of ICA is The Blind Signal Separation (BSS). In signal separation, multiple streams of information are extracted from linear mixtures of these signal streams. This process is blind if the examples of the source signals, along with their corresponding mixtures, are unavailable for training. In BSS, signals are estimated from their unknown linear mixtures with the assumption that the sources are mutually independent. In recent times, there has been considerable interest in various neural realizations of ICA and BSS. In these approaches, the higher order statistics are typically taken into account by using suitable non-linearities in the learning phase, even though the final output mapping is still linear. In the statistical approaches, separation is usually achieved by optimizing some constraint functions that are defined explicitly in terms of the cumulants (higher order statistics) of the observed data. On the other hand, in

---

Corresponding Author: B. Justus Rabi, College of Engineering, Anna University, Chennai-25

neural approaches to BSS and ICA, cumulants are replaced by suitable non-linearities in the learning algorithms. The non-linearities implicitly introduce the higher order statistics which are necessary for blind separation.

In ICA[5] the data vectors are represented in a linear basis which is determined by requiring that the coefficients of expansion must be mutually independent. Therefore the basis vectors of ICA are generally nonorthogonal and higher order statistics are needed in determining the ICA expansion. However this kind of representation often characterizes the fundamental properties of the data better than standard PCA. For example in blind signal separation of the original source signals.

The ICA operates on M zero mean source signals $s_k(1)....,S_K(m)$, k=1, 2,…, that are scalar-valued and mutually statistically independent for each sample value k. The original sources are unobservable and the input to the ICA are different linear mixtures $x_k(1),....,x_k(L)$ of the sources. The signal model in ICA is written in vector form as:

$$M_{xk} = A * s_k = \Sigma \; s_k(i)*a(i) \qquad (1) \qquad i=1$$

Here $s_k = [S_K(1),…, S_K(M)]^T$ is the source vector consisting of the M source signals $S_K(I)$ (i=1,…, M) at the index value k., A=[a(1),….,A(M)] is a constant LxM mixing matrix whose elements are the unknown coefficients of the mixtures. The columns a (I) are the basis vectors of ICA. Usually M=L i.e. the number of source M is assumed to be equal to the number of available different mixtures L, to simplify the derivation of BSS algorithms. In practice, it is not necessary for M to be equal to L. Furthermore, each source signals $S_K(1)$ is a stationary zero mean stochastic process. Only one of the source signals $S_K(I)$ is allowed to have a Gaussian distribution. This follows from the fact that it is impossible to separate several Gaussian's from each other.

The very little prior information is assumed on the matrix A. Therefore the stronger independence assumptions are required in determining the ICA expansion. Even then only the directions of ICA basis vectors a (I), I=1, 2,3, …………. M are defined, because their magnitudes and amplitude of the source signals $S_K(I)$ can be interchanged in the model (1). Also the order of the terms in the sum in (1) can be arbitrary. To get a more Unique expansion, one can either require that each source $S_K(I)$ has unit variance or normalize the basis vectors a (I) to unit length (and then arranged for them according to the powers of the sources).

Linear models of the form (1) are used in several known techniques, but the assumptions are different .In the standard least square method, that the matrix A is completely known. That it is easy to estimate the vector $s_k s_k = (A^T A)^{-1}$ at X_k. If the matrix A is known except for a few parameters, subspace type methods or the minimum like hood method can be used for estimating

the unknown parameters. In standard PCA, the expansion (1) is determined by requiring that the basis vectors a (i) are orthonormal and the coefficients $S_K(I)$ have maximal variant.

Blind signal separation: Blind Source separation has lately become a popular research area both in statistical signal processing and unsupervised neural learning. In BSS the goal is to separate mutually statistically independent but otherwise unknown source signals from their linear mixtures without knowing the mixing coefficients.

The abilities, generalization and self organization of artificial neural network have motivated novel studies in applying it to the feature extraction and blind separation. ANN's, biologically motivated and statistically based models, have been used successfully for highly parallel and error tolerant machines.

Blind Source Separation and the closely related independent Component Analysis (ICA) have been studied in statistical signal processing since 1980's. In these statistical approaches, separation is usually achieved by optimizing some constraint functions that are defined explicitly in terms of the cumulants of the observed data. In BSS[6] the objective is to separate mutually statistically independent but otherwise known source signals from their linear mixtures without knowing the mixing coefficients.

The task is to find {sk (i)} of the sources, with only the data vectors x_k and the number of sources M known.

The whitening process is required for the input vectors so that (i) the data vectors x_k has a zero mean (ii) the variance of the observed signals are normalized to unity (iii) separation algorithms have better stability properties and converge faster (iv) The components of the whitened vectors v_k are mutually uncorrelated since it is a necessary prerequisite for the stronger independence condition. If the data vectors have non-zero mean they first subtract from them. Furthermore the effects of higher order statistics can be removed by using the whitening transformation. Standard Principal Component Analysis is often used for whitening, since one can then simultaneously compress information optimally in the mean square error sense and filter possible noise. However Whitening the data can make the separation problem more difficult if the mixing matrix is ill-conditioned or if some of the source signals are relatively weak compared to other signals. The input vectors x (k) is whitened by applying transformation v (k) =Vx (k). Where v(k) is the kth whitened vector and V is the whitening matrix. The whitening matrix can be determined in two ways: by using a batch and the approach or by neural learning. For the batch approach, if PCA is used to determine the whitening matrix is given by $V=D^{-1/2} E^T$ where the matrix V is chosen so that the covariance matrix $v_k v_k^T$ equals the unit matrix I_M. The matrix D =diag[$\lambda(1),…..,$ $\lambda(M)$] is the MxM diagonal matrix, where $\lambda(i)$ is the ith largest eigenvalue of the data covariance matrix $E\{x_k x_k^T\}$. The matrix E =[c(1),…,c(M)] is an LxM

matrix, where c[i] denotes the respective i[th] principal eigenvector. Therefore the transformation actually consists of two steps that is compressed and whitening. The compression step consists of selecting the proper value for q (the number of source signals). Therefore the PCA described above for whitening can also be used to select (i.e.,) estimate the number of source signals q to be recovered.(or the number of independent components)if the noise term n(k) is assumed to be zero-mean Gaussian white noise with covariance matrix. Probably the simplest neural algorithm for learning the whitening matrix Vk is:

$$V_{k+1} = V_k - u_k(V_k V_k^T - 1)\ V_k \qquad (2)$$

Whitening and related procedures have been some criticized, because they do not provide uniform performance in subsequent operation. In uniform performance methods, the separation capability does not depend on the condition number of the mixing matrix A. In theory, it is then possible to separate even very weak sources or use almost similar mixtures as inputs. However this property assumes that the input data obeys the ICA model exactly with no noise. In our experiments, separation algorithms that require prewhitening usually performed quite well in normal conditions, for example when the mixing matrix A was chosen randomly.

The separation process can be carried out using many different methods. Separation can be achieved by optimizing some constraint functions that are defined explicitly in terms of the cumulants of the observed data. However contrast functions typically require extensive batch computations using estimated higher order statistics of the data and lead to very complicated adaptive separation algorithms. As we will see it is sufficient to use the Kurtosis (fourth order cumulant) of the data. There is an alternative to the neural learning approach presented here that converges much faster. Another class of separation methods involves using neural networks to perform the separation of the source signals. In fig. Neural network model the second stage of architecture is responsible for the separation of whitened signals v. The linear separation transformation is given by:

$$y(k) = W^T v(k) \qquad (3)$$

$W^T W = I_q$ is the separation matrix thus the separated signals are the outputs of the second stage, that is s(k0=y(k). An interesting observation is that once the source signals (k) have been estimated, this means the pseudo inverse of A is that A$^+$, must have been blindly determined.

Nonlinear pica subspace learning rule: This algorithm was originally introduced by Oja some years ago as an extension of his well known PCA subspace rule In BSS it is used quite similar but the update formula is different.

$$W_{k+1} = W_k + \mu_k\ [\ v_k - W_k\ g\ (y_k)]\ g(y_k^T) \qquad (4)$$

Here the learning parameter μ_k must be positive for stability reasons, restricting the applicability of mainly to sub- Gaussian sources. Without prewhitening, nonlinear PCA subspace rule is able to somehow separate sinusoidal type sources only, because the algorithm responds in this case still largely to the second order statistics.

A major advantage of this learning rule is that can be realized using a simple modification of one layer standard symmetric PCA network, allowing a simple and local neutral implementation. An interesting feature is that the underlying data model is actually slightly nonlinear. Non-linear PCA subspace rule can then be derived by approximately minimizing the mean square error. In spite of this, the algorithm performs well for sub-Gaussian sources even in larger problems. Also for good convergence, it is best to select the initial weight matrix W (0) to have as columns a set of orthonormal vectors. Typically the nonlinear function is chosen as:

$$(t) = \beta\ \tanh\ (t/\beta) \qquad (5)$$

Where $g(t) = df(t)/dt$ and $f(t) = \beta^2 \ln\ [\cosh(t/\beta)]$. This is not an arbitrary choice for the non linearity in the learning rule of (4). It is motivated by the fact that when determining the ICA expansion higher –order statistics are needed. This can be seen by observing another neural learning rule to perform the separation of unknown signals. This learning rule called Bigradient algorithm given by:

$$W(k+1) = W(k) + v(k)\mu_k\ g(y_k^T)$$
$$+ \gamma\ (k)\ \gamma\ (k)\ W(k)\ [1 - W^T\ (k)\ W(k)] \qquad (6)$$

where, $\gamma\ (k)$ is another gain parameter, typically about 0.5 or 1. This is a stochastic gradient algorithm that maximizes or minimizes the performance criterion:

$$J(W) = \sum E\ \{\ f(y_i)\ \} \qquad (7)$$

Under the constraint that the weight matrix W is orthonormal. The orthonormal constraint in (7) is realized in the learning rule (6) in an additive manner. With the appropriate function in (7), the performance criterion would involve the sum of the fourth order statistics (fourth order cumulants) of the outputs that is the Kurtosis. Therefore the criterion would be either minimized for sources with a negative Kurtosis or maximized for the sources with a positive kurtosis .Source signals that have a positive Kurtosis are referred to as Super Gaussian signals. In (7) the expectation operator would be dropped because we only consider instantaneous values.

The logistic function f(t)=ln [cosh(t)] (for β =1) in terms of a Taylor's series expansion:

$$f(t) = \ln\ [\cosh(t)] = t^2\ /2\ -t^4/4\ + t^6/45\ - \qquad (8)$$

The second order term $t^2/2$ is on the average constant due to the whitening. The non linearity would then be given by $g(t) = df(t)/dt = t - t^3/3 + 2t^6/15 - \ldots$ and the cubic term will be dominating (an odd function) if the data are prewhitened.

Estimation process[7] is the last stage in the neural network model. Two methods are presented here to estimate the ICA basis vectors, or the column vectors of the mixing matrix A. The first method is a batch approach where the estimate of A that is Â is given by:
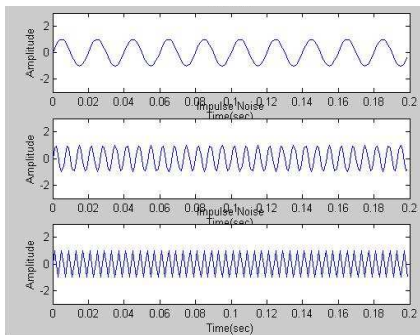
$$\hat{A} = ED^{1/2}W \ldots (10)$$

E has columns that are the associated eigenvectors and W is the separation matrix. The second method is a neural approach for estimating the ICA basis vectors. From the neural network model gives an estimate of the observed data:
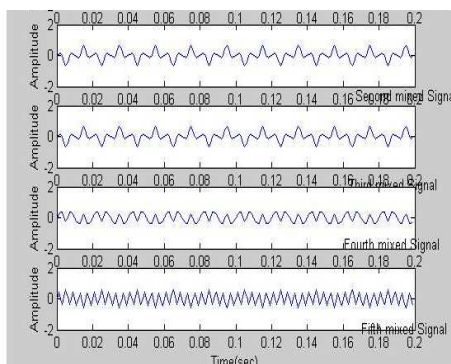
$$X = Q y \qquad (11)$$

Comparing (11) with (1) for n=0 (i.e ., x=As), we see that $Q = \hat{A}$ since y =s . Therefore the columns of Q matrix are estimates of the columns of A, the ICA basis vectors. A neural learning algorithm can be derived from a representation error performance measure given by:

$$J(Q) = \tfrac{1}{2} \|X - \overline{X}\|^2_2 = \tfrac{1}{2} \| x - Qy \|^2_2 \qquad (12)$$
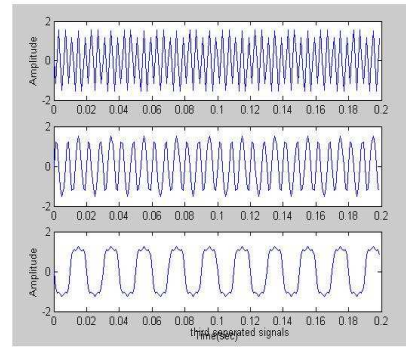
SIMULATION RESULTS

Input signals for training the network

Sampled mixed signals

Separated signals from the neural network

CONCLUSION

The harmonics present in the inverters is eliminated using Blind Signal Separation. This algorithm eliminates any type of harmonic mixture present in the signal. This is carried out in the simulation using Matlab Software. Implementation of the algorithm for estimating the harmonics is achieved by hardware. In future the set of algorithm can be implemented in a neural chip (single IC), thereby reducing the space and cost.

REFERENCES

1. Karhunen, J. and P. Pajunen, 1997. Blind source separation using at least square type adaptive algorithms. In: Proc. IEEE Intl. Conf. On Acoustics, Speech and Signal Processing (ICASSP'97) Munch, Germany, pp: 3361-3364.
2. Amari, S., A. Chicoki and H. Yang, 1996. A New learning algorithm for Blind signal Separation. To Appear in Advances Neural Information Processing Systems (8) (Proc NIPS '95), Cambridge Ma: MIT Press.
3. Karhunen, J., 1996. Neural approaches to independent component analysis and source separation. In: Proc. 4th European Symp. On Artificial Neural Network (ESSANN'96) Brugesa, Belgium, pp: 249-266.
4. Mark, P., 2001. Conditions for non-negative independent component analysis. IEEE Signal Processing Letters, Jul. 17.
5. Heinz, M., M. Joho and G.S. Moschytz, 2000. A simple threshold non linearity for blind separation of sub-Gaussian signals. IEEE Intl. Symp. On Circuits and Systems ISCAS 2000, May 28-31, Geneva, Switzerland, pp: iv 489-492.
6. Henz, M., T.P. Von Hoff and M. Joho, 2002. Blind separation of signals with mixed kurtosis signs using threshold activation functions. IEEE Trans. On Neural Networks, 12: 3.
7. Karhunen, J., E. Oja, L. Wang, R. Vigario and J. Joutesensalo, 1997. A class of neural networks for independent component analysis. IEEE Trans. on Neural Networks, 8: 486-504.