Batch Mutually Referenced Separation Algorithm for MIMO Convolutive Mixtures

A. Mansour

Lab. E^3I^2, ENSIETA,
29806 Brest cedex 09, (FRANCE).
mansour@ieee.org
http://www.ensieta.fr
http://ali.mansour.free.fr

Abstract. This paper deals with the blind separation problem of Multi-Input Multi-Output (MIMO) convolutive mixtures. Previously, we presented some algorithms based on mutually referenced criterion to separate MIMO convolutive mixtures. However, the proposed algorithms are time consuming and they need a lot of computation efforts. It is obvious that the computation efforts can be reduced as well the convergence time when the adaptive algorithms are well initialized. To choose the best starting point of these algorithms, we propose here a direct and batch minimization of the proposed criteria.

1 Introduction

The blind separation of sources (BSS) (or the Independent Component Analysis "ICA") problem consists of the estimation of the unknown input signals of an unknown channel using only the output signals of that channel (i.e., the observed signals or the mixing signals) [1, 2]. The sources are assumed to be statistically independent from each other [3]. Recently, that problem has been addressed and applied in many different situations [4] such as speech enhancement [5], separation of seismic signals [6], sources separation method applied to nuclear reactor monitoring [7], airport surveillance [8], noise removal from biomedical signals [9], and some radar applications have been addressed in [10].

Generally, the transmission channel is considered to be a memoryless channel (i.e., the case of an instantaneous mixture) or a matrix of linear filters (i.e. in the convolutive mixture case). Since 1985, many algorithms have been proposed to solve the ICA problem [11–15]. The criteria of those algorithms have generally been based on high-order statistics [16–18]. Recently, by using only second-order statistics, some subspace methods that blindly separate the sources in the case of convolutive mixtures have been explored [19, 20].

In previous work, we proposed two subspace approaches using LMS [20, 21] or a conjugate gradient algorithm [22] to minimize subspace criteria. Those criteria were derived from the generalization of the method proposed by Gesbert *et*
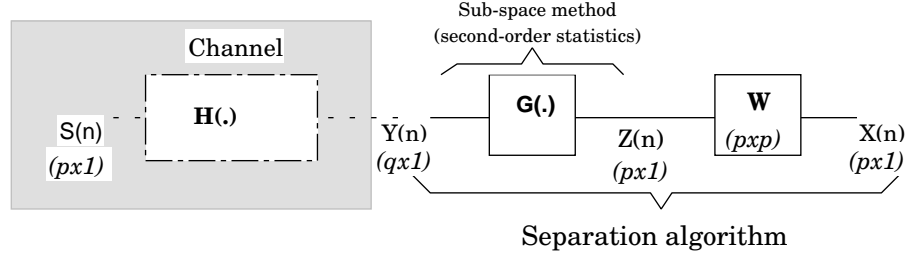
Fig. 1. General Structure.

al. [23] for blind identification[1]. The main advantage of such algorithms is that one can identify almost all parameters of the mixing filters using only second order statistics. The major drawbacks of the previous proposed algorithms are computation efforts and time consuming algorithms. In fact, the convergence of these algorithms is very slow and it can be improved by better choice of algorithm initialization parameters. In the following we discuss some initialization scenarios.

2   Channel Model, Assumptions and Background

Let $S(n)$ denotes the p unknown sources which are statistically independent from each other. $Y(n)$ is the $q \times 1$ observed vector, see Fig. 1. The relationship between $S(n)$ and $Y(n)$ is given by :

$$Y(n) = [\mathcal{H}(z)]S(n) \tag{1}$$

where $\mathcal{H}(z)$ stands for the channel effect. In the case of convolutive mixture, $\mathbf{H}(z) = (h_{ij}(z))$ becomes a $q \times p$ complex polynomial matrix. In the following, we consider that the channel is a linear and causal one and that the coefficients $h_{ij}(z)$ are RIF filter. Let M denotes the degree of the channel which is the highest degree of $h_{ij}(z)$. The previous equation (1) can be rewritten as:

$$Y(n) = \sum_{i=0}^{M} \mathbf{H}(i)S(n-i) \tag{2}$$

Here $\mathbf{H}(i)$ denotes the $q \times p$ real constant matrix corresponding to the impulse response of the channel at time i and $S(n-i)$ is the $p \times 1$ source vector at time $(n-i)$. Considering $(N+1)$ observations of the mixture vector $(N > q)$ and

---

[1] In the identification problem, the authors generally assume that there is only one source and that the source is an independent and identically distributed (iid) signal.

using the following notations:

$$Y_N(n) = \begin{pmatrix} Y(n) \\ \vdots \\ Y(n-N) \end{pmatrix} \quad \text{and} \quad S_{M+N}(n) = \begin{pmatrix} S(n) \\ \vdots \\ S(n-M-N) \end{pmatrix}, \quad (3)$$

model (2) can be rewritten as:

$$Y_N(n) = \mathbf{T}_N(\mathbf{H})S_{M+N}(n), \quad (4)$$

where the $q(N+1) \times p(M+N+1)$ matrix $\mathbf{T}_N(\mathcal{H})$ is the Sylvester matrix corresponding to $\mathcal{H}(z)$. In reference [24], the Sylvester matrix is given by:

$$\mathbf{T}_N(\mathbf{H}) = \begin{bmatrix} \mathbf{H}(0) & \mathbf{H}(1) & \mathbf{H}(2) & \cdots & \mathbf{H}(M) & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{H}(0) & \mathbf{H}(1) & \cdots & \mathbf{H}(M-1) & \mathbf{H}(M) & \mathbf{0} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \cdots & \cdots & \mathbf{0} & \mathbf{H}(0) & \mathbf{H}(1) & \cdots & \mathbf{H}(M) \end{bmatrix}. \quad (5)$$

Under some mild and realistic assumptions [19] (i.e the number of sensors is larger than the number of sources $q > p$ and H(z) is a column-reduced irreducible matrix), we proposed in [20] a subspace separation algorithm based on the identification algorithm proposed by Gesbert *et al.* in [23, 25]. That algorithm consists of two steps:

 – A blind identification step is done by minimizing a second order criterion

$$\mathcal{C}(\mathrm{G}) = \mathrm{E} \, \|(\mathbf{I} \quad \mathbf{0})\mathbf{G}Y_N(n) - (\mathbf{0} \quad \mathbf{I})\mathbf{G}Y_N(n+1)\|^2. \quad (6)$$

Here, E stands for the expectation, $\mathbf{G}$ denotes a $p(M+N+1) \times q(N+1)$ real matrix and $\mathbf{I}$ is the $(M+N)p \times (M+N)p$ identity matrix. It has been shown in [20] that the above minimization leads to a matrix $\mathbf{G}^\star$ such:

$$\mathbf{Perf} = \mathbf{G}^\star \, \mathbf{T}_N(\mathbf{H}) = \mathrm{diag}(\mathbf{A}, \cdots, \mathbf{A}), \quad (7)$$

where **Perf** denotes the performance matrix and $\mathbf{A}$ is any $p \times p$ matrix.
 – It is clear that the first step does not achieve the separation due to fact that the residual matrix $\mathbf{A}$ isn't a general permutation matrix [3]. Therefore, one should apply any high order statistics BSS algorithm of instantaneous mixtures (we used different algorithms as [13, 26]).

Later on, we derived another algorithm based on the same criterion [27]. In [27], we proved the type and the uniqueness of solution (7). We should mention that equation (7) can lead us to an spurious solution $\mathbf{G}$. To avoid such solution, the minimization of equation (6) has be done with respect to a consistent constraint, please see [27]:

$$\mathbf{G}_1\mathbf{R}_Y(n)\mathbf{G}_1^T = \mathbf{I}_p, \quad (8)$$

where $\mathbf{G}_i$ is the ith block row $p \times q(N+1)$ of $\mathbf{G}$, $\mathbf{R}_Y(n) = EY_N(n)Y_N(n)^T$ is the covariance matrix of $Y_N(n)$ and $\mathbf{I}_p$ is a $p \times p$ identity matrix. If the above constraint is satisfied and $\mathbf{G_1}$ is such that $\mathbf{G}_1 Y_N(n) = \mathbf{A}S(n)$, then:

$$\mathbf{G}_1 \mathbf{R}_Y(n) \mathbf{G}_1^T = \mathbf{A}\mathbf{R}_S(n)\mathbf{A}^T = \mathbf{I}_p, \tag{9}$$

where $\mathbf{R}_S(n) = ES(n)S(n)^T$ is the source covariance matrix. $\mathbf{R}_S(n)$ is a full rank diagonal matrix as a result of the statistical independence of the p sources from each other. When equation (9) is satisfied, matrix $\mathbf{A}$ becomes invertible.

3   Batch Algorithm

To improve the performance and the convergence speed of the previous proposed algorithms [20, 27], one can at first simplify the constraint. Actually, the proposed constraint (8) is equivalent to p equations. That constraint can be easily change to one equation constraint such as:

$$||\mathbf{G}_1 \mathbf{R}_Y(n) \mathbf{G}_1^T|| = 1 \tag{10}$$

Another constraint can also be derived as:

$$\det\{\mathbf{G}_1 \mathbf{R}_Y(n) \mathbf{G}_1^T\} = 1 \tag{11}$$

One should mention that the last two equations (10) and (11) can avoid spurious solutions as well as done the first constraint. In addition, the constant "1" used in both equations can be change to any other positive number since that will be reduced to a simple normalization of the residual matrix $\mathbf{A}$.

3.1   Criterion Minimization

Using the criterion derivative found in [20], we can prove that the minimization of the criterion is equivalent to the following matrix equation system:

$$\mathbf{G}_1 \mathbf{R}_Y(n) = \mathbf{G}_2 \mathbf{R}_Y^T(n+1)$$
$$2\mathbf{G}_2 \mathbf{R}_Y(n) = \mathbf{G}_3 \mathbf{R}_Y^T(n+1) + \mathbf{G}_1 \mathbf{R}_Y(n+1)$$
$$\vdots$$
$$2\mathbf{G}_i \mathbf{R}_Y(n) = \mathbf{G}_{i+1} \mathbf{R}_Y^T(n+1) + \mathbf{G}_{i-1} \mathbf{R}_Y(n+1) \tag{12}$$
$$\vdots$$
$$\mathbf{G}_{M+N+1} \mathbf{R}_Y(n) = \mathbf{G}_{M+N} \mathbf{R}_Y^T(n+1)$$

Here $\mathbf{R}_Y(n) = EY_N(n)Y_N(n+1)^T$ is the correlation matrix of $Y_N(n)$ and $Y_N(n+1)$.

Under the channel assumptions considered in the previous section, Sylvester matrix $\mathbf{T}_N(\mathbf{H})$ becomes a full rank matrix [27]. Using the previous statement, the

fact that the sources are supposed to be persistently exciting and the definition of $\mathbf{R}_Y(n)$ one can easily prove that $\text{Rank}(\mathbf{R}_Y(n)) = (M + N + 1)p$. Hence, one can find using a SVD decomposition two matrices $\mathbf{U}$ and $\mathbf{V}$ such that $\mathbf{U}$ is a $q(N+1) \times (M+N+1)p$ left invertible matrix and $\mathbf{V}$ is a $(M+N+1)p \times q(N+1)$ right invertible matrix. Let $\mathbf{V}^\dagger$ be the right pseudo-inverse matrix of $\mathbf{V}$ and let us denote $\mathbf{P} = \mathbf{R}_Y^T(n+1)\mathbf{V}^\dagger$ and $\mathbf{Q} = \mathbf{R}_Y(n+1)\mathbf{V}^\dagger$, then system (12) becomes:

$$\mathbf{G}_1\mathbf{U} = \mathbf{G}_2\mathbf{P}$$
$$2\mathbf{G}_2\mathbf{U} = \mathbf{G}_3\mathbf{P} + \mathbf{G}_1\mathbf{Q}$$
$$\vdots$$
$$2\mathbf{G}_i\mathbf{U} = \mathbf{G}_{i+1}\mathbf{P} + \mathbf{G}_{i-1}\mathbf{Q} \tag{13}$$
$$\vdots$$
$$\mathbf{G}_{M+N+1}\mathbf{U} = \mathbf{G}_{M+N}\mathbf{Q}$$

To minimize the criterion, one can solve the above matrix equation system (13).

3.2   Analytical Solution

In this subsection, an analytical solution of system (13) is given. Using the fact that $\mathbf{U}$ is a full rank matrix and a QR decomposition [28], one can find an orthogonal matrix $\mathbf{L}$ and an upper triangular matrix $\mathbf{R}$ such that $\mathbf{U} = \mathbf{LR}$. System (13) can then be rewritten as

$$\tilde{\mathbf{G}}_1\mathbf{R} = \tilde{\mathbf{G}}_2\tilde{\mathbf{P}}$$
$$2\tilde{\mathbf{G}}_2\mathbf{R} = \tilde{\mathbf{G}}_3\tilde{\mathbf{P}} + \tilde{\mathbf{G}}_1\tilde{\mathbf{Q}}$$
$$\vdots$$
$$2\tilde{\mathbf{G}}_i\mathbf{R} = \tilde{\mathbf{G}}_{i+1}\tilde{\mathbf{P}} + \tilde{\mathbf{G}}_{i-1}\tilde{\mathbf{Q}} \tag{14}$$
$$\vdots$$
$$\tilde{\mathbf{G}}_{M+N+1}\mathbf{R} = \tilde{\mathbf{G}}_{M+N}\tilde{\mathbf{Q}}$$

Where $\tilde{\mathbf{G}}_i = \mathbf{G}_i\mathbf{L}^T$, $\tilde{\mathbf{P}} = \mathbf{P}^T\mathbf{P}$ and $\tilde{\mathbf{Q}} = \mathbf{Q}^T\mathbf{Q}$. Since $\mathbf{R}$ is a full rank upper triangular matrix, then without loss of generality one can write

$$\mathbf{R} = \begin{pmatrix} \bar{\mathbf{R}} \\ \mathbf{0} \end{pmatrix} \tag{15}$$

where $\mathbf{0}$ is a zero matrix of appropriate dimensions. Let us decompose the three matrices $\tilde{\mathbf{P}}$, $\tilde{\mathbf{Q}}$ and $\tilde{\mathbf{G}}_i$ as following

$$\tilde{\mathbf{P}} = \begin{pmatrix} \tilde{\mathbf{P}}_1 \\ \tilde{\mathbf{P}}_2 \end{pmatrix}$$
$$\tilde{\mathbf{Q}} = \begin{pmatrix} \tilde{\mathbf{Q}}_1 \\ \tilde{\mathbf{Q}}_2 \end{pmatrix} \tag{16}$$
$$\tilde{\mathbf{G}}_i = \begin{pmatrix} \tilde{\mathbf{g}}_i & \tilde{\mathbf{X}}_i \end{pmatrix}$$

where the different sub-matrices are of appropriate dimensions. Using equations (14) and (16), one can write

$$\tilde{\mathbf{g}}_1 = \tilde{\mathbf{g}}_2\bar{\mathbf{P}}_1 + \tilde{\mathbf{X}}_2\bar{\mathbf{P}}_2$$
$$2\tilde{\mathbf{g}}_2\mathbf{R} = \tilde{\mathbf{g}}_3\bar{\mathbf{P}}_1 + \tilde{\mathbf{X}}_3\bar{\mathbf{P}}_2 + \tilde{\mathbf{g}}_1\bar{\mathbf{Q}}_1 + \tilde{\mathbf{X}}_1\bar{\mathbf{Q}}_2$$
$$\vdots$$
$$2\tilde{\mathbf{g}}_i\mathbf{R} = \tilde{\mathbf{g}}_{i+1}\bar{\mathbf{P}}_1 + \tilde{\mathbf{X}}_{i+1}\bar{\mathbf{P}}_2 + \tilde{\mathbf{g}}_{i-1}\bar{\mathbf{Q}}_1 + \tilde{\mathbf{X}}_{i-1}\bar{\mathbf{Q}}_2 \qquad (17)$$
$$\vdots$$
$$\tilde{\mathbf{g}}_{M+N+1}\mathbf{R} = \tilde{\mathbf{g}}_{M+N}\bar{\mathbf{Q}}_1 + \tilde{\mathbf{g}}_{M+N}\bar{\mathbf{Q}}_2$$

where $\bar{\mathbf{P}}_i\bar{\mathbf{R}} = \tilde{\mathbf{P}}_i$ and $\bar{\mathbf{Q}}_i\bar{\mathbf{Q}} = \tilde{\mathbf{Q}}_i$.

Theoretically, any solution of the previous system (17) can minimize the proposed criterion. We should mention here that system (17) contents $M+N+1$ matrix equations and $2(M+N+1)$ unknown matrices (i.e. $\tilde{\mathbf{g}}_i$ and $\tilde{\mathbf{X}}_i$) which means that we have many solutions. These solutions are natural solutions (up to permutation and scale polynomial filter) and spurious solution as mentioned in the previous section.

3.3   Simplified Approximation

In [29], $\mathbf{R}_Y(n)$ is considered to be a full rank matrix (i.e the channel is noisy one), then system (12) with respect of the constraint can be be solved as following

$$\mathbf{D}_{(i+1)} = \mathbf{B}(2\mathbf{I} - \mathbf{D}_i\mathbf{A})^{-1} \qquad (18)$$
$$\mathbf{G}_{M+N-i-1} = \mathbf{G}_{M+N-i-2}\mathbf{D}_i \qquad (19)$$

where $\mathbf{D}_0 = \mathbf{B} = \mathbf{R}_N(n+1)\mathbf{R}_N^{-1}(n)$ and $\mathbf{A} = \mathbf{R}_N^T(n+1)\mathbf{R}_N^{-1}(n)$. It is proved that with a Signal to Noise Ratio (SNR) over 15dB, the previous system (19) can give satisfactory results.

Due to limitation of page number, Another simplified approximation of he system (17) has been omitted.

4   Conclusion

In this paper, a batch mutually algorithm for MIMO convolutive mixtures is presented. Generally, the batch algorithms have several advantages over adaptive algorithms from computation efforts and time point of view. However, they are very sensitive to noisy channel and to the estimation errors of the different needed parameters. Therefore, they are promising solutions to solve the initialization problems of adaptive algorithms and to improve also their performances.

References

1. A. Hyvärinen and E. Oja, "Independent componenet analysis: algorithms and applications," *Neural Networks*, vol. 13, pp. 411–430, 2000.

2. A. Mansour, A. Kardec Barros, and N. Ohnishi, "Blind separation of sources: Methods, assumptions and applications.," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E83-A, no. 8, pp. 1498–1512, August 2000.

3. P. Comon, "Independent component analysis, a new concept?," *Signal Processing*, vol. 36, no. 3, pp. 287–314, April 1994.

4. A. Mansour and M. Kawamoto, "Ica papers classified according to their applications & performances.," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E86-A, no. 3, pp. 620–633, March 2003.

5. L. Nguyen Thi and C. Jutten, "Blind sources separation for convolutive mixtures," *Signal Processing*, vol. 45, no. 2, pp. 209–229, 1995.

6. N. Thirion, J. Mars, and J. L. Boelle, "Separation of seismic signals: A new concept based on a blind algorithm," in *Signal Processing VIII, Theories and Applications (EUSIPCO'96)*, Triest, Italy, September 1996, pp. 85–88, Elsevier.

7. G. D'urso and L. Cai, "Sources separation method applied to reactor monitoring," in *Proc. 3rd Workshop on Higher Order Statistics*, Edinburgh, Scotland, September 1994.

8. E. Chaumette, P. Comon, and D. Muller, "Application of ICA to airport surveillance," in *HOS 93*, South Lake Tahoe-California, 7-9 June 1993, pp. 210–214.

9. A. Kardec Barros, A. Mansour, and N. Ohnishi, "Removing artifacts from ECG signals using independent components analysis," *NeuroComputing*, vol. 22, pp. 173–186, 1999.

10. M. Bouzaien and A. Mansour, "HOS criteria & ICA algorithms applied to radar detection," in *4th International Workshop on Independent Component Analysis and blind Signal Separation, ICA2003*, Nara, Japan, 1-4 April 2003.

11. J. Hérault, C. Jutten, and B. Ans, "Détection de grandeurs primitives dans un message composite par une architecture de calcul neuromimétique en apprentissage non supervisé," in *Actes du Xème colloque GRETSI*, Nice, France, 20-24, May 1985, pp. 1017–1022.

12. J. F. Cardoso and P. Comon, "Tensor-based independent component analysis," in *Signal Processing V, Theories and Applications (EUSIPCO'90)*, L. Torres, E. Masgrau, and M. A. Lagunas, Eds., Barcelona, Espain, 1990, pp. 673–676, Elsevier.

13. C. Jutten and J. Hérault, "Blind separation of sources, Part I: An adaptive algorithm based on a neuromimetic architecture," *Signal Processing*, vol. 24, no. 1, pp. 1–10, 1991.

14. S. I. Amari, A. Cichocki, and H. H. Yang, "A new learning algorithm for blind signal separation," in *Neural Information Processing System 8*, Eds. D.S. Toureyzky et. al., 1995, pp. 757–763.

15. A. Mansour and C. Jutten, "A direct solution for blind separation of sources," *IEEE Trans. on Signal Processing*, vol. 44, no. 3, pp. 746–748, March 1996.

16. M. Gaeta and J. L. Lacoume, "Sources separation without a priori knowledge: the maximum likelihood solution," in *Signal Processing V, Theories and Applications (EUSIPCO'90)*, L. Torres, E. Masgrau, and M. A. Lagunas, Eds., Barcelona, Espain, 1990, pp. 621–624, Elsevier.

17. N. Delfosse and P. Loubaton, "Adaptive blind separation of independent sources: A deflation approach," *Signal Processing*, vol. 45, no. 1, pp. 59–83, July 1995.

18. A. Mansour and C. Jutten, "Fourth order criteria for blind separation of sources," *IEEE Trans. on Signal Processing*, vol. 43, no. 8, pp. 2022–2025, August 1995.

19. A. Gorokhov and Ph. Loubaton, "Subspace based techniques for second order blind separation of convolutive mixtures with temporally correlated sources," *IEEE Trans. on Circuits and Systems*, vol. 44, pp. 813–820, September 1997.

20. A. Mansour, C. Jutten, and Ph. Loubaton, "An adaptive subspace algorithm for blind separation of independent sources in convolutive mixture," *IEEE Trans. on Signal Processing*, vol. 48, no. 2, pp. 583–586, February 2000.

21. A. Mansour, C. Jutten, and Ph. Loubaton, "Subspace method for blind separation of sources and for a convolutive mixture model," in *Signal Processing VIII, Theories and Applications (EUSIPCO'96)*, Triest, Italy, September 1996, pp. 2081–2084, Elsevier.

22. A. Mansour, A. Kardec Barros, and N. Ohnishi, "Subspace adaptive algorithm for blind separation of convolutive mixtures by conjugate gradient method," in *The First International Conference and Exhibition Digital Signal Processing (DSP'98)*, Moscow, Russia, June 30-July 3 1998, pp. I–252–I–260.

23. D. Gesbert, P. Duhamel, and S. Mayrargue, "Subspace-based adaptive algorithms for the blind equalization of multichannel fir filters," in *Signal Processing VII, Theories and Applications (EUSIPCO'94)*, M.J.J. Holt, C.F.N. Cowan, P.M. Grant, and W.A. Sandham, Eds., Edinburgh, Scotland, September 1994, pp. 712–715, Elsevier.

24. T. Kailath, *Linear systems*, Prentice Hall, 1980.

25. D. Gesbert, P. Duhamel, and S. Mayrargue, "On-line blind multichannel equalization based on mutually referenced filters," *IEEE Trans. on Signal Processing*, vol. 45, no. 9, pp. 2307–2317, September 1997.

26. A. Mansour and N. Ohnishi, "Multichannel blind separation of sources algorithm based on cross-cumulant and the levenberg-marquardt method.," *IEEE Trans. on Signal Processing*, vol. 47, no. 11, pp. 3172–3175, November 1999.

27. A. Mansour, "A mutually referenced blind multiuser separation of convolutive mixture algorithm," *Signal Processing*, vol. 81, no. 11, pp. 2253–2266, November 2001.

28. G. H. Golub and C. F. Van Loan, *Matrix computations*, The johns hopkins press-London, 1984.

29. A. Mansour and N. Ohnishi, "A batch subspace ica algorithm.," in *10th IEEE Signal Processing Workshop on Statistical Signal and Array Processing*, Pocono Manor Inn, Pennsylvania, USA, 14 - 16 August 2000, pp. 63–67.