HAWKES-LAGUERRE REDUCED RANK MODEL FOR POINT PROCESSES

Syed Ahmed Pasha

School of Electrical & Information Engineering
University of Sydney
NSW 2006 Australia

Victor Solo

School of Electrical Engineering & Telecom.
University of New South Wales
NSW 2052 Australia

ABSTRACT

In recent years there has been a surge in the demand for analysis tools for multivariate point process data driven by work in neural coding and high frequency finance. In both these areas data volumes have become huge but few dimension reduction methods have been developed. Here we introduce a reduced rank model for the multivariate point process and provide a maximum likelihood estimator which we compute by an NMF type algorithm. However, the dependence on the point process history in the model implies our algorithm does not fit the traditional framework. The method is illustrated with a simulation and some data from cortical recordings from cats.

Index Terms— Point process, stochastic intensity, reduced rank, NMF, maximum likelihood

1. INTRODUCTION

The availability of high-dimensional multichannel spike recordings in neuroscience [1, 2] and more recently tick-level data in stochastic finance [3] have led to a surge in demand for analysis tools for multivariate point processes.

The breakthrough in point process theory with the development of stochastic intensity [4, 5] in the late 1960s led to considerable development of dynamic point process models [6, 7] with applications in the recent neural coding literature (see e.g. [8, 9] and references therein). But these methods do not address the curse of dimensionality problem and become unwieldy for high-dimensional point processes.

This is traditionally overcome by dimension reduction strategies such as principal components analysis (PCA) [10, 11] but until recently [12] there was no true PCA for multivariate point processes; one which did not temporally bin the point process which loses temporal information [2].

An important technique for dimension reduction in the time series literature is the reduced rank regression discussed in [13] which imposes a rank restriction on the matrices. A point process analog can be constructed for the mutually exciting or multivariate Hawkes process [14] which additionally imposes non-negativity constraints on the matrices.

Recently, popularized by [15, 16], a suite of algorithms have been proposed for non-negative matrix factorization (NMF) which have found immediate appeal in a number of application areas e.g. [17, 18, 19]. In the traditional NMF framework there is no notion of a dynamic model so the fitting algorithm under non-negativity constraints for the point process system which depends on the point process history cannot be cast in the traditional framework. Here we derive a NMF-type algorithm for the dynamic point process model for the first time. The relation between our algorithm and other models for NMF in the literature is further elucidated in section 3.

Relation to Prior Work: We make some comments on the relation between this paper and [12, 20]. [12] introduced a principled approach to dimension reduction for multivariate point processes via a PCA model as well as dynamic index model (DIM). Positivity of the stochastic intensity was ensured by developing models for the log-stochastic intensity. However, there is no history dependence in the PCA models of [12, 20] and the DIM in [12] was introduced in the log-stochastic intensity model which has been found to be sensitive to variations in starting values for the iterative optimization and furthermore difficult to interpret. Here however we propose a Hawkes-Laguerre reduced rank (HL-RR) model which depends on the point process history and the factorization into non-negative basis and components matrices provides a natural interpretation.

In the remainder of the paper we introduce the HL-RR model in section 2. We develop a steepest descent algorithm (section 3) via a cyclic descent procedure for maximum likelihood. The algorithm is demonstrated on a simple simulation example (section 4.1) and then tested on neuronal recordings in cat primary visual cortex (section 4.2). The paper concludes with some final remarks in section 5.

2. HAWKES-LAGUERRE REDUCED RANK MODEL

We observe a d-dimensional multivariate point process N_τ consisting of counting processes $N_{k,\tau}, k = 1, \cdots, d$. Here $N_{k,\tau} = \#$ events of the k-th process up to time and including τ. Assuming No-Simultaneity [12] or orderliness [5] (i.e. in a small time interval only one event of any type can occur) we

can define the vector stochastic intensity,

$$\mu_{k,\tau} = P(N_{k,\tau+\delta} - N_{k,\tau} = 1|\mathcal{H}^\tau) = \mu_{k,\tau}\delta + o(\delta), k = 1, ..., d$$

where $\mathcal{H}^\tau$ denotes the past of the vector process $N_s, 0 < s \leq \tau$. This is often written informally as $\mu_\tau d\tau = E(dN_\tau|\mathcal{H}^\tau)$.

Then Jacod's multivariate log-likelihood [5] is given by,

$$\mathcal{L} = \Sigma_1^d \int_0^T (ln\mu_{k,\tau} dN_{k,\tau} - \mu_{k,\tau} d\tau)$$

In our previous work [12, 20] we introduced the canonical parameter $\theta_{k,\tau} = \ln\mu_{k,\tau}$ so that the log-likelihood is revealed as a member of the exponential family and obtained dimension reduction in the canonical parameter. Here we construct a Hawkes-Laguerre reduced rank (HL-RR) model but for the stochastic intensity.

$$\begin{aligned} \mu_\tau &= \bar{\mu} + \int_0^\tau H(u) dN_{\tau-u} \\ &= \bar{\mu} + \int_0^\tau F_{d\times q} G_{q\times d}(u) dN_{\tau-u} \\ &= \bar{\mu} + F \int_0^\tau G(u) dN_{\tau-u} \end{aligned}$$

where $\bar{\mu} = (\bar{\mu}_1, \ldots, \bar{\mu}_d)^T$ is the d-vector of unknown background firing rates and we have $d > q$. We now expand $G(u)$ in Laguerre polynomials. This is a causal basis and helps preserve positivity.

$$G(u) = \Sigma_1^p G_l (u\beta)^{l-1} e^{-\beta u}$$

where $\frac{1}{\beta}$ is a user chosen time constant. The entries of the $q \times d$ matrices $G_l = [g_{c,j;l}]$ for $l = 1, ..., p$ will need to be non-negative. We also denote $F = [f_{k,c}]$ and require $f_{k,c} > 0$. Continuing

$$\begin{aligned} \int_0^\tau G(u) dN_{\tau-u} &= \Sigma_1^p G_l \int_0^\tau (u\beta)^{l-1} e^{-\beta u} dN_{\tau-u} \\ &= \Sigma_1^p G_l \psi_{\tau;l} \end{aligned}$$

Note that $\psi_{\tau;l}$ can be precomputed and so can be assumed known. We thus have

$$\begin{aligned} \mu_\tau &= \bar{\mu} + F\Sigma_1^p G_l \psi_{\tau;l} \\ \equiv \mu_{k,\tau} &= \bar{\mu}_k + \Sigma_1^p f_k^T G_l \psi_{\tau;l} \\ &= \bar{\mu}_k + \Sigma_{l=1}^p \Sigma_{c=1}^q \Sigma_{j=1}^d f_{k,c} g_{c,j;l} \psi_{j;\tau;l} \quad (1) \end{aligned}$$

Remark: The model depends on the product FG_l. There is a potential identifiability issue since $FG_l = FMM^{-1}G_l$ where M is a scaled permutation [21]. To remove this lack of identifiability we must place constraints on F, G_l but then the non-negativity of F, G_l is not preserved. In the traditional non-negative matrix factorization (NMF) algorithm [15] the scaling ambiguity is dealt with by normalizing the columns of F but this is not the case here.

Notwithstanding the identifiability issue, there has been a growing interest in NMF which provides physical interpretation of latent structure [15, 19, 17]. Here we expect to gain similar insight.

3. MAXIMUM LIKELIHOOD VIA CYCLIC DESCENT

We develop an NMF-type algorithm for the point process HL-RR model. However our algorithm does not fit in the traditional NMF framework. Firstly (1) does not give an NMF type decomposition of the mean since it depends on the point process history. Further, the unknown parameters are not two matrices as in traditional NMF but rather $p + 1$ matrices, namely F and $G_1, \cdots, G_p$. Even if $p = 1$ we still do not have a traditional setup because of the presence of the data dependent $\psi_{j;\tau;l}$ terms in (1).

We also note that again the presence of the data dependent factors $\psi_{j;\tau;l}$ means that our model is not the same as the convolutive model [22, 23, 24]. Finally our model is not the same as the non-negative PARFAC model [25, 26, 27] again because of the presence of the known data dependent term $\psi_{j;\tau;l}$.

Nevertheless we are able to derive an NMF-type algorithm in section 3.1.

It is known [28] that the traditional Kullback-Liebler type updates form in fact an expectation-maximization (EM) algorithm for a Poison type model. It would be of interest to see if such an interpretation is possible for the algorithm we now develop; this will be pursued elsewhere.

We partition the interval $0 < \tau \leq T$ into tiny bins of width δ so that $N_{k\tau}^\delta = N_{k,\tau+\delta} - N_{k,\tau}$ is 0 or 1 with very high probability. Let $T = n\delta$ and $\tau = t\delta, t = 1, .., n$, then

$$\begin{aligned} \mathcal{L} &\sim \Sigma_1^d \Sigma_1^n [N_{kt}^\delta ln\mu_{k,t} - \mu_{k,t}\delta] \\ \Rightarrow -\frac{1}{\delta}\mathcal{L} &\sim \Sigma_1^d \Sigma_1^n [-Y_{kt} ln\mu_{k,t} + \mu_{k,t}] \quad (2) \end{aligned}$$

where $Y_{kt} = \frac{1}{\delta} N_{kt}^\delta$.

In section 3.1 by an abuse of notation we use $\psi_{j;t;l}$ instead of $\psi_{j;t\delta;l}$ and write $G_{1,p}$ to mean $G_1, ..., G_p$.

3.1. Multiplicative Updates

We use the following Majorization-Minimization (MM) update result from [16]. Consider the negative log-likelihood (2),

$$J = \Sigma_i [-Y_i ln\mu_i + \mu_i]$$

where $\mu_i = \bar{\mu}_i + \Sigma_m W_{im}\theta_m$ and $\theta_m \geq 0, W_{im} \geq 0, Y_i \geq 0$.

Then [16] J is decreased by the multiplicative update

$$\theta_m^{(1)} = \frac{\theta_m^{(0)}}{\Sigma_i W_{im}} \Sigma_i W_{im} \frac{Y_i}{\mu_i^{(0)}}, \quad \mu_i^{(0)} = \bar{\mu}_i + \Sigma_m W_{im} \theta_m^{(0)}$$

The trick in using this result is to realize that we can take i and m to be multi-indices.

We can modify J by adding terms that depend only on Y_i. Thus minimizing J is equivalent to minimizing

$$D = \Sigma_i D_i = \Sigma_i [Y_i ln \frac{Y_i}{\mu_i} - Y_i + \mu_i]$$

The inequality $xlnx - x + 1 \geq 0$ rapidly delivers $D_i \geq 0$ (divide through by μ_i). D_i is the Kullback-Liebler distance between Y_i and μ_i. It is D that [16] deals with but the equivalence to likelihood is well known.

Now we pursue a three stage cyclic descent minimization [29],

$$F\text{-step: Given } G_{1,p}, \bar{\mu} \text{ get } F^* = \text{arg.min}_{\substack{F:f_{k,c}\geq 0, \\ F^T\mathbf{1}=\mathbf{1}}} D$$

$$G\text{-step: Given } F, \bar{\mu} \quad \text{get } G_l^* = \text{arg.min}_{G_l:g_{c,j;l}\geq 0} D$$

$$\bar{\mu}\text{-step: Given } F, G_{1,p} \text{ get } \bar{\mu}^* = \text{arg.min}_{\bar{\mu}:\bar{\mu}\geq 0} D$$

The $F, G, \bar{\mu}$ steps of the cyclic descent minimization are constrained optimizations placing non-negativity constraints on the entries. We have found that the multiplicative MM algorithm provides a reliable approach to this cyclic ascent.

F-step. The criterion is separable in k so we can take k fixed. If we identify $m \equiv c$ and $i \equiv t$ then we have the equivalences

$$\begin{aligned} &\mu_i \equiv \mu_{k,t}, \bar{\mu}_i \equiv \bar{\mu}_{k,t} = \bar{\mu}_k, \text{ and } Y_i \equiv Y_{kt} \\ &\theta_m \equiv f_{k,c} \end{aligned}$$

and then

$$\begin{aligned} \mu_{k,t} &= \bar{\mu}_k + \Sigma_{c,j,l} f_{k,c} g_{c,j;l} \psi_{j;t;l} \\ &= \bar{\mu}_k + \Sigma_c W^f_{tc} f_{k,c} \\ W^f_{tc} &= \Sigma_{j,l} g_{c,j;l} \psi_{j;t;l} \end{aligned}$$

This delivers the updates for $k = 1, \cdots d$

$$\begin{aligned} f^{(1)}_{k,c} &= \frac{f^{(0)}_{k,c}}{\Sigma_t W^f_{tc} + \lambda_c} \Sigma_t W^f_{tc} \frac{Y_{kt}}{\mu^{(0)}_{k,t}} \\ \mu^{(0)}_{k,t} &= \bar{\mu}^{(0)}_k + \Sigma_c W^f_{tc} f^{(0)}_{k,c} \\ W^f_{tc} &= \Sigma_{j,l} g^{(0)}_{c,j;l} \psi_{j;t;l} \end{aligned}$$

where λ_c is the Lagrange multiplier due to the constraint $\Sigma_k f_{k,c} = 1$.

G-step. Here we identify

$$i \equiv (k,t) \text{ and } m \equiv (c,j,l)$$

Then we have

$$\begin{aligned} &Y_i \equiv Y_{kt}, \mu_i \equiv \mu_{k,t} \text{ and } \bar{\mu}_i \equiv \bar{\mu}_{k,t} = \bar{\mu}_k \\ &\theta_m \equiv g_{c,j;l} \end{aligned}$$

and then

$$\begin{aligned} \mu_{k,t} &= \bar{\mu}_k + \Sigma_{c,j,l} g_{c,j;l} f_{k,c} \psi_{j;t;l} \\ &= \bar{\mu}_k + \Sigma_{c,j,l} W^g_{ktcjl} g_{c,j;l} \\ W^g_{ktcjl} &= f_{k,c} \psi_{j;t;l} \end{aligned}$$

This delivers the update

$$\begin{aligned} g^{(1)}_{c,j;l} &= \frac{g^{(0)}_{c,j;l}}{\Sigma_{kt} W^g_{ktcjl}} \Sigma_{kt} W^g_{ktcjl} \frac{Y_{kt}}{\mu^{(0)}_{k,t}} \\ \mu^{(0)}_{k,t} &= \bar{\mu}^{(0)}_k + \Sigma_{c,j,l} g^{(0)}_{c,j;l} f^{(1)}_{k,c} \psi_{j;t;l} \\ W^g_{ktcjl} &= f^{(1)}_{k,c} \psi_{j;t;l} \end{aligned}$$

$\bar{\mu}$-step. The steepest descent in the $\bar{\mu}$-step is similar to the F-step in that the criterion is separable in k so we can take k fixed. We rapidly find the updates for $k = 1, ..., d$

$$\begin{aligned} \bar{\mu}^{(1)}_k &= \frac{\bar{\mu}^{(0)}_k}{n} \Sigma_t \frac{Y_{kt}}{\mu^{(0)}_{k,t}} \\ \mu^{(0)}_{k,t} &= \bar{\mu}^{(0)}_k + \Sigma_{c,j,l} f^{(1)}_{k,c} g^{(1)}_{c,j;l} \psi_{j;\tau;l} \end{aligned}$$

Note that wherever in the sums $Y_{kt} = 0$ then the corresponding term is omitted from the sum.

4. SIMULATION AND DATA ANALYSIS

A simulation example is first provided to demonstrate the algorithm. Then, data analysis of multichannel recordings from cortical neurons in cats is presented.

4.1. Simulation Example

A $d = 5$-dimensional Hawkes process [14] is considered. The simulation setup is discussed first. We consider the simple case with $p = 1, q = 2$ and generate matrices F, G_1 whose entries are non-negative. The background firing rate $\bar{\mu}_k, k = 1, ..., d \sim \mathcal{U}(4, 5)$ and the time constant $\frac{1}{\beta} = \frac{2}{3}$ s is taken. The multivariate Hawkes process is simulated in $0 < \tau \leq T$ with $T = 10$ s by extending the algorithm for the bivariate case [30] based on the thinning procedure [31]. About 70 counts per channel were recorded and at the resolution of $\delta = 0.1$ ms the No-Simultaneity condition [12] was ensured.

Fig. 1(a),(b) show the components $G_1 \psi_{\tau;1}$ and the rows of F^T respectively. Looking at the individual weights in the first row we see that channel 2 gets a weighting of about 0.5 and channel 3 gets a weighting of about 0.3. Thus the first component is roughly a weighted average over these channels. For the second row we find roughly a weighted average of channels $1, 4, 5$.

We use the Akaike Information Criterion (AIC) for model comparison, computed using the likelihood since it only needs μ_τ. As shown in Fig. 1(c), the minimizer of the AIC is the true value of $q = 2$.

For the cyclic descent minimization, the initial estimates of F, G_1 were randomly generated with non-negative entries. The Kullback-Leibler (KL) divergence measure $\Sigma_i D_i(Y_i || \mu_i)$ for the true parameter values and the iterates computed using the parameter estimates are shown in Fig. 1(d). We observe

that the iterates converge after about 40 iterations to a smaller value than the true one. We do not show iterates of individual F, G_1 parameters. Even for a problem of small dimension such a plot is dense and difficult to interpret. Instead we show lower dimensional summary measures; namely column cosines which are the angles between a true column and the corresponding estimate of it. Fig. 1(e) shows the column cosines diag($\hat{F}^T F$) of F where $\hat{F}$ denotes the estimate of F. Similarly, Fig. 1(f) shows the column cosines of (normalized) G_1^T; some bias is evident. Fig. 1(g),(h) show the components for the estimate of G_1^T and the rows of $\hat{F}^T$ respectively. Notice that the components as well as the rows of the $\hat{F}^T$ are interchanged. We have found that the results are reasonably robust to variations in starting values for the iterative optimization.

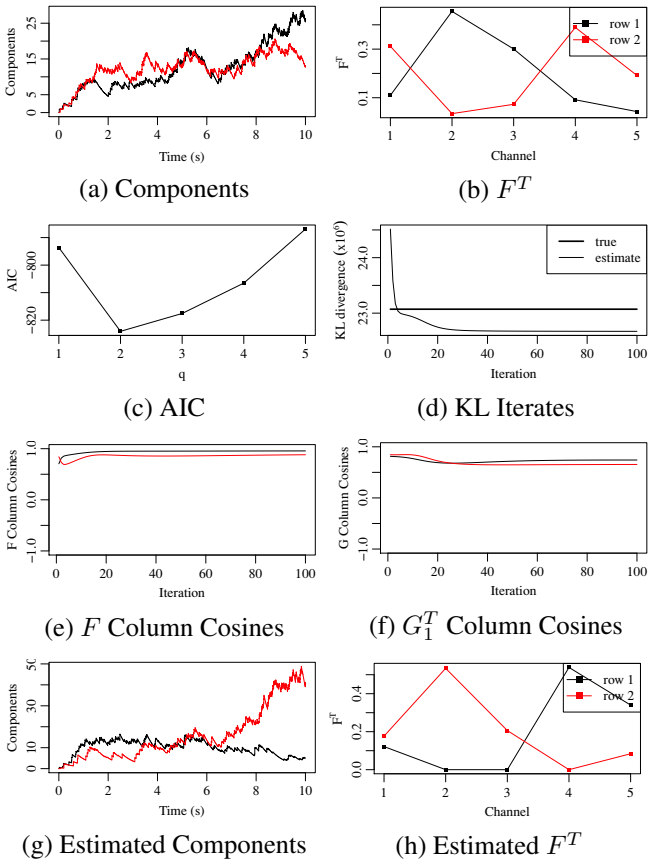

Fig. 1. Simulated Data: (a) Components $G_1\psi_{1t}$, (b) Rows of F^T, (c) Akaike Information Criterion (AIC) Values, (d) Iterates of KL Divergence measure, (e) Iterates of F Column Cosines, (f) Iterates of G_1^T Column Cosines, (g) Estimated Components and (h) Estimated F^T.

4.2. Neural Data

Extracellular neuronal recordings in the cat primary visual cortex are analyzed. Up to 25 neurons (channels) were recorded simultaneously for about 2.5 min from area 17 (trans-columnar down the medial bank) [32]. For illustration purposes we present results for the analysis of spiking activity for 4 s in $d = 9$ channels with about 100 spikes per channel. The discretization step $\delta = 0.08$ ms and $\frac{1}{\beta} = 0.5$ s were taken.

$q = 2$ was determined as the minimizer of the AIC shown in Fig. 2(a). For $q = 2$, the iterates of the KL divergence are shown in Fig. 2(b) which settle after about 70 iterations.

Fig. 2(c),(d) show the estimated components and basis vectors respectively. We find that for the first row channel $6, 8$ get a weighting of about $0.3, 0.25$ respectively. Thus the first component is roughly a weighted average over these channels. A weighted average of channels $2, 5, 7, 8$ contributes to the second row.

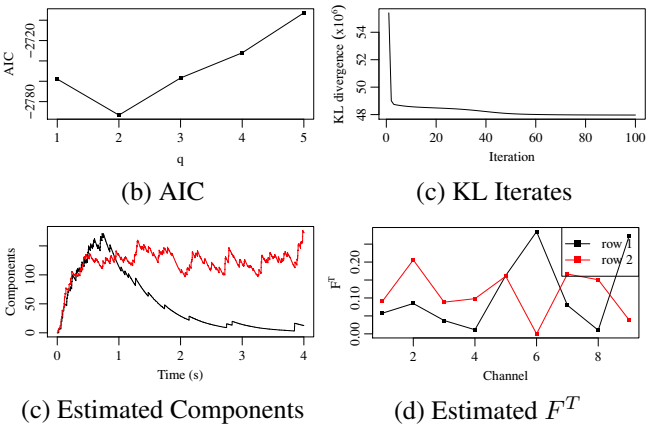

Fig. 2. Neural Data: (a) Akaike Information Criterion (AIC) Values, (b) Iterates of KL Divergence measure, (c) Estimated Components and (d) Estimated Basis Vectors.

5. CONCLUSIONS

In this paper we have developed a multivariate point process reduced rank (RR) model which we estimated by maximum likelihood and fitted by a novel NMF type algorithm. The RR model preserves positivity of the intensity by placing non-negativity constraints on the matrices. The presence of a dynamic point process model implies that our fitting algorithm is considerably different to existing NMF algorithms. The approach is illustrated with a simulation example and tested on neural recordings.

Acknowledgement. Neural data were recorded by Tim Blanche in the laboratory of Nicholas Swindale, University of British Columbia, and downloaded from the NSF-funded CRCNS Data Sharing website.

6. REFERENCES

[1] P. Dayan and L.F. Abbott, *Theoretical Neuroscience*, MIT Press, Cambridge MA, 2001.

[2] F. Rieke, D. Warland, R. de Ruyter van Stevenink, and W. Bialek, *Spikes: Exploring the Neural Code*, MIT Press, Boston, 1997.

[3] N. Hautsch, *Modelling Irregularly Spaced Financial Data*, Springer, New York, 2004.

[4] D.L. Snyder, *Random Point Processes*, J. Wiley, New Jersey, 1975.

[5] D.J. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes, Volume I (2nd. ed.)*, Springer-Verlag, New York, 2003.

[6] D.R. Brillinger, "The identification of point process systems," *Ann. Prob.*, vol. 3, pp. 909–929, 1975.

[7] D.R. Brillinger, "Maximum likelihood analysis of spike trains of interacting nerve cells," *Biol. Cybern.*, vol. 59, pp. 189–200, 1988.

[8] S. Kim et. al., "A Granger causality measure for point process models of ensemble neural spiking activity," *PLoS Comput Biol*, vol. 7(3), 2011.

[9] R.E. Kass, R.C. Kelly, and W.L. Loh, "Assessment of synchrony in multiple neural spike trains using loglinear point process models," *Ann. Appl. Stat.*, vol. 5, 2011.

[10] I.T. Jolliffe, *Principal Components Analysis*, Springer Verlag, Heidelberg, 1986.

[11] D.R. Brillinger, *Time Series: data analysis and theory*, Holden-Day, Inc., San Francisco, 1981.

[12] V. Solo, "High dimensional point process system identification: PCA and dynamic index models," in *Proc. IEEE Conf. Decision & Control*, 2006, pp. 829–833.

[13] G.C. Reinsel and R.P. Velu, *Multivariate Reduced Rank Regression*, Springer, Berlin, 1998.

[14] A.G. Hawkes, "Point spectra of some mutually exciting point processes," *J. Roy. Statist. Soc.*, vol. 33(3), 1971.

[15] D.D. Lee and H.S. Seung, "Learning the parts of objects by nonnegative matrix factorization," *Nat*, vol. 401, 1999.

[16] D.D. Lee and H.S. Seung, "Algorithms for non-negative matrix factorization," in *NIPS*, 2000.

[17] M.W. Berry et. al., "Algorithms and applications for approximate nonnegative matrix factorization," in *Comput. Stat. Data An.*, 2006, pp. 155–173.

[18] M.D. Plumbley, A. Cichocki, and R. Bro, "Non-negative mixtures," in *Handbook of Blind Source Separation: Independent Component Analysis and Applications*, pp. 515–547. Academic Press, 2010.

[19] J.P. Brunet, P. Tamayo, T.R. Golub, and J.P. Mesirov, "Metagenes and molecular pattern discovery using matrix factorization," *PNAS*, vol. 101(12), 2000.

[20] V. Solo and S.A. Pasha, "Point process principal components analysis via geometric optimization," *Neural Computation*, vol. 25(1), pp. 101–122, 2013.

[21] H. Minc, *Nonnegative Matrices*, John Wiley & Sons, New York, NY, USA, 1988.

[22] P. Smaragdis, "Non-negative matrix factor deconvolution; extraction of multiple sound sources from monophonic inputs," in *Independent Component Analysis and Blind Signal Separation*, vol. 3195 of *Lec Notes Comput Sci*, pp. 494–499. 2004.

[23] P. Smaragdis, "Convolutive speech bases and their application to supervised speech separation," *IEEE Audio, Speech, Language Process*, vol. 15(1), pp. 1–12, 2007.

[24] T. Virtanen, "Separation of sound sources by convolutive sparse coding," in *Proc. SAPA*, 2004.

[25] R. Bro, "PARAFAC. Tutorial and applications," *Chemom. Intell. Lab. Syst.*, vol. 38(2), 1997.

[26] J.D. Carroll, G.D. Soete, and S. Pruzansky, "Multiway data analysis," in *Fitting of the Latent Class model via iteratively reweighted least squares CANDECOMP with nonnegativity constraints*, pp. 463–472. 1989.

[27] W.P. Krijnen and J.M.F.T. Berge, "Contrastvrije oplossingen van het CANDECOMP/PARAFAC-model," *Kwantitatieve Methoden*, vol. 12, pp. 87–96, 1991.

[28] A.T. Cemgil, "Bayesian inference for nonnegative matrix factorisation models," *Comput Intell Neurosci*, 2009.

[29] D.G. Luenberger, *Linear and non-linear programming*, J. Wiley, New York, 1974.

[30] M. Ogata, "On Lewis' simulation method for point processes," *IEEE Trans. Inf. Theory*, vol. 27(1), pp. 22–31, 1981.

[31] P.A.W. Lewis and G.S. Shedler, "Simulation of nonhomogeneous Poisson processes by thinning," *Naval Research Logistics*, vol. 26, pp. 403–413, 1979.

[32] T.J. Blanche et. al., "Polytrodes: high density silicon electrode arrays for large scale multiunit recording," *J. Neurophys.*, vol. 93(5), pp. 2987–3000, 2005.