ON A PRACTICAL APPROACH TO SOURCE SEPARATION OVER FINITE FIELDS FOR NETWORK CODING APPLICATIONS

Irina Nemoianu[1], Claudio Greco[1], Marc Castella[2], Béatrice Pesquet-Popescu[1], Marco Cagnazzo[1]

[1]*TELECOM-ParisTech, TSI department, 46 rue Barrault, 75634 Paris Cedex 13, FRANCE*

[2]*Télécom & Management SudParis, CITI departement, 9 rue Charles Fourier, 91011 Evry Cedex, FRANCE*

{nemoianu,greco,cagnazzo,pesquet}@telecom-paristech.fr,marc.castella@it-sudparis.eu

ABSTRACT

In Blind Source Separation, or BSS, a set of source signals are recovered from a set of mixed observations without knowledge of the mixing parameters. Originated for real signals, BSS has recently been applied to finite fields, enabling more practical applications. However, classical entropy-based techniques do not perform well in finite fields. Here, we propose a non-linear encoding of the sources to increase the discriminating power of the separation methods. Our results show that the encoding improves the success rate of the separation for sources with few samples in large finite fields, both conditions met in practical networking applications. Our results open new possibilities in the context of network coding –wherein linear combinations of packets are sent in order to maximize throughput and increase loss immunity– by relieving the nodes from the need to send the combination coefficients, thus reducing the overhead cost.

Index Terms— Blind Source Separation, Channel Coding, Galois Fields, Independent Component Analysis, Network Coding.

1 Introduction

Blind Source Separation (BSS) [1, 2] consists in recovering a set of source signals $\mathbf{S}$ from a set of mixed signals $\mathbf{X} = f(\mathbf{S})$, also referred to as *observations*, without knowing the sources themselves nor the mixing process parameters. This is a subject that has been intensively investigated in the last three decades, due to its potential numerous applications in speech recognition, sensor/biomedical signal processing, *etc*.

The *Independent Component Analysis* (ICA) [3, 4] approach solves the BSS problem relying on the assumption that the sources are statistically independent and non-Gaussian. Given a set of observations, ICA algorithms return a set of estimated source signals that maximize a separation criterion, referred to as *contrast function*. Separation criteria can be based on information-theoretic principles, *e.g.*, maximizing the entropy or minimizing a Kullback-Leibler divergence, while other approaches build on higher order statistics. In any case, the assumptions of independence and non-Gaussianity are explicitly used.

One should note that ICA can only retrieve the original sources up to some ambiguities: there will be a permutation ambiguity, *i.e.*, the algorithm will not be able to tell which reconstructed source is which, and scaling ambiguity, *i.e.*, the reconstructed sources will be identified up to a scaling factor.

ICA has been recently extended to the case of finite fields [5], which presents several additional challenges due to the nature of the operations defined over a finite field. In particular, a technique can be based on the fact that the entropy of any linear combination of statistically independent random variables over GF(q) is larger than the entropy of any of the components, as long as none of them is uniform. Separation is therefore possible by finding the inverse linear transformation that minimizes the marginal entropy of the resulting combinations. Since the operations take place in a finite field, an exhaustive approach is possible, *i.e.*, to try any possible linear combinations of observations until we find the one that has the lowest entropy [5]. While the method was introduced at first as an interesting theoretical result, its potential can be seen for practical applications too. For instance, it has been suggested that BSS schemes over finite fields can be used in the context of eavesdropping over MIMO multi-user digital communications systems [6].

Another very interesting potential application for an efficient source separation algorithm over finite fields is in the design of a transmission scheme similar to Network Coding (NC) [7]. In NC, instead of merely relaying packets, the intermediate nodes of a network send linear combinations of the packets they have previously received, with random coefficients taken from a finite field [8–11]. NC, used as an alternative to traditional routing, has proved beneficial to real-time streaming applications, both in terms of maximization of the throughput and in terms of reduction of the effects of losses [12–17]. However, in practical Network Coding approaches, the random coefficients must be added to the packet as headers [11], incurring an overhead that can be prohibitive if the maximum packet size is small. On the other hand, in a BSS based approach, it could be possible to relieve the nodes from the need to include the coefficients in the packets, thus reducing significantly the amount of data that has to be transmitted to the receiver in order to decode the packets. Such an approach would instead rely on the capability of the receivers to reconstruct the coefficient themselves.

In this article we improve the results of the separation method by increasing the discriminating power of the algorithm without adding constraints on the distribution of the sources. The rationale is that many of the sources in today's applications do have a distribution close to the uniform, *e.g.*, compressed videos images or sound, so the above methods fail in this case. We propose to pre-process the sources with a non-linear encoding which, as we will show, increases the separability of the ICA method towards a more practical application with higher GF orders and sources closer to uniform.

The rest of this article is organized as follows: in Section 2 we give an overview on some relevant related work. In Section 3, we introduce our proposed approach for blind source separation for sources in GF(2^b) and the rationale behind it. Then in Section 4 we validate our approach with experimental results and a comparison with a state-of-the-art exhaustive entropy-based source separation algorithm. In Section 5 we draw conclusions and outline future work. Finally, in the Appendix, we quantify the augmented discriminating power granted to the algorithm by the non-linear coding we introduced.

2 Related Work

Several algorithms have been proposed to reduce the search space and the execution time of blind source separation algorithms, at the expenses of the accuracy [18, 19].

One such technique has been proposed for finite fields of prime order, but can be easily extended to the general case [5]. At each iteration, the algorithm finds a couple of observation vectors $\mathbf{x}_i$ and $\mathbf{x}_j$ and a scalar k in the finite field such that $H(\mathbf{x}_i+k\mathbf{x}_j)<H(\mathbf{x}_i)$ and replaces $H(\mathbf{x}_i)$ with $H(\mathbf{x}_i+k\mathbf{x}_j)$. When no possible substitution can be found, the algorithm terminates, and the final value of the $\mathbf{x}_i$ will be the reconstruction of the original sources. This algorithm is significantly faster then an exhaustive search, but is prone to local minima. Other methods have been proposed, *e.g.*, approximating the entropy with $-p\log(p)$, where p is the probability of the most probable element [18, 19].

Since the scope of this paper is focused on success rate rather than complexity, we shall compare ourself to the Ascending Minimization of Entropies for ICA method [5], originally proposed for GF(2). This method extracts a single source, then removes the contribution from this source to the mixtures and repeats this process N times, after which it has found all N sources, restricting the search space to vectors linearly independent from the ones recovered so far. Our technique will also follow the same approach, but the search space will be further restricted to vectors that yield admissible sources, *i.e.*, codewords.

Algorithm 1 Separation algorithm.

Input: $(N \times T)$ mixed sample matrix $\mathbf{X}$.
Output: $(N \times T)$ separated source matrix $\widetilde{\mathbf{S}}$.
$\mathcal{V} \leftarrow \emptyset$, $\mathcal{W} \leftarrow \emptyset$;
for all $\mathbf{w}$ of length N in GF $\left(2^{b+1}\right)$ **do**
  $\widetilde{\mathbf{z}} \leftarrow \mathbf{w}^\top \mathbf{X}$;
  if $\widetilde{\mathbf{z}}$ is a codeword **then**
    $\mathcal{V} \leftarrow \mathcal{V} \cup \{\mathbf{w}\}$;
  end if
end for
repeat
  $\mathbf{w}^* \leftarrow \arg\min_{\mathbf{w}\in\mathcal{V}} \left\{H\left(\textsc{decode}\left(\mathbf{w}^\top \mathbf{X}\right)\right)\right\}$;
  if $\mathbf{w}^* \notin \textsc{span}(\mathcal{W})$ **then**
    $\mathcal{W} \leftarrow \mathcal{W} \cup \{\mathbf{w}^*\}$;
  end if
  $\mathcal{V} \leftarrow \mathcal{V} - \{\mathbf{w}^*\}$;
until $\|\mathcal{W}\| = N$
$\mathbf{W} \leftarrow$ matrix built from the row vectors in $\mathcal{W}$;
$\widetilde{\mathbf{Z}} \leftarrow \mathbf{W}^\top \mathbf{X}$;
$\widetilde{\mathbf{S}} \leftarrow \textsc{decode}\left(\widetilde{\mathbf{Z}}\right)$;

3 Proposed approach

In this section, we describe our proposed method to separate a number of linearly combined (*mixed*) independent sources defined in a finite field. Generally speaking, the ability of an algorithm to identify a source given a set of mixed observation (*demixing*) stems from the ability to identify a property that holds true for the original sources and does not for the mixtures. For instance, entropy based methods assume that the original sources have lower entropy than the mixtures.

Our main idea is to increase the discriminating power of the algorithm by pre-processing the sources with an error-detecting code. The code should be such that the probability of a mixture belonging to the code is small. Also, the code cannot be linear, otherwise mixtures would always belong to it; we therefore consider only non-linear codes.

A simple example of non-linear code is the odd-parity bit-code. A parity bit-code is a systematic code consisting in adding a *parity bit* to the source symbol to ensure that the number of bits with the value one in the encoded symbol is always even (even-parity bit-code) or odd (odd-parity bit-code). Parity bit codes are the simplest form of error detecting code, and have been in use, both in hardware and in software applications, since the 1950s. For our purposes, we use an odd-parity bit-code because it is obviously non-linear, as the null-string is not a codeword (since it has zero bits with value one and zero is an even number). A detailed analysis of the discriminating power of the odd-parity bit-code, *i.e.*, its ability to distinguish between sources and mixtures, is given in the Appendix.

Let us now consider a set of N independent source signals $\mathbf{s}_0, \mathbf{s}_1, \ldots, \mathbf{s}_{N-1}$, each containing T samples, defined in a finite field GF $\left(2^b\right)$. First of all, the sources are encoded with an odd-parity bit-code, such that each element in the encoded source belongs to GF $\left(2^{b+1}\right)$, because of the added parity bit, and has an odd number of bits equal to one in its binary representation. Let us call $\mathbf{z}_n$ the encoded version of a source

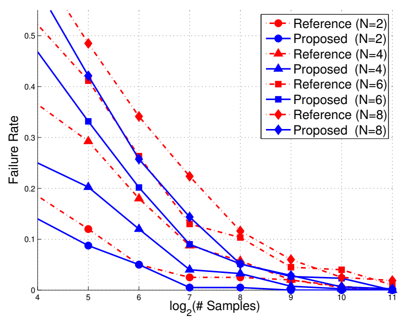

Fig. 1. Comparison between the reference and the proposed technique for finite field GF(2). The failure rate, *i.e.*, the percentage of sources that the algorithm was not able to identify, is plotted against the number of samples in the mixture in log-scale.

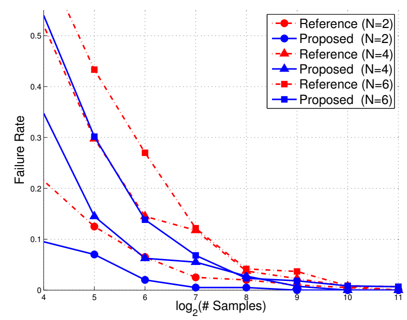

Fig. 2. Comparison between the reference method and the proposed technique for finite field GF(4). The failure rate is plotted against the number of samples in the mixture in log-scale.

$\mathbf{s}_n$, and $\mathbf{Z}$ the N-by-T matrix which has $\mathbf{z}_n$ as its n-th row, for $n \in \{0 \ldots N-1\}$. These encoded sources are combined with an unknown N-by-N mixing matrix $\mathbf{A}$, also defined in GF $\left(2^{b+1}\right)$: $\mathbf{X}{=}\mathbf{A}\mathbf{Z}$.

In order for our separation problem to have a solution, we assume that the matrix $\mathbf{A}$ is invertible, *i.e.*, rank$(\mathbf{A}){=}N$. Each row $\mathbf{x}_n$ of $\mathbf{X}$ is a linear combination, or mixture, of the encoded sources. In order to recover the original sources, we proceed according to Algorithm 1, as follows. For each vector $\mathbf{w}$ of length N in GF $\left(2^{b+1}\right)$, we try to *demix* one encoded source $\widetilde{\mathbf{z}}{=}\mathbf{w}^\top\mathbf{X}$.

If all T elements of $\widetilde{\mathbf{z}}$ are codewords, we decode the vector, *i.e.*, we remove the parity bit from its elements, thus obtaining $\widetilde{\mathbf{s}}$, and estimate the entropy $H(\widetilde{\mathbf{s}})$. Notice that the probability of a random mixture being a codeword decreases with T. After all the vectors $\mathbf{w}$ have been tried, we select the N linearly independent vectors corresponding to the demixed sources with the lowest entropy. The matrix $\mathbf{W}$ composed as the horizontal concatenation of these vectors is our estimation of the inverse matrix of $\mathbf{A}$. We limit ourselves to a family of linearly independent vectors under the assumption that, being $\mathbf{W}$ the inverse of $\mathbf{A}$, it has full rank N. The demixed sources corresponding to this matrix $\widetilde{\mathbf{Z}} = \mathbf{W}^\top\mathbf{X}$ will represent our estimation of the encoded sources. It will suffice to remove the parity bits in order to recover the original sources up to a scaling and permutation ambiguity.

4 Experimental Results

In the following, we present the results relative to the separation of N sources of T elements for the proposed bit-code based technique, and compare them with the results achievable using an exhaustive entropy-based technique at the same rate. The reference technique simply consists in identifying the N linear combinations of observations such that the combination coefficients are linearly independent and the entropy is minimized [5, 19]. Our technique, on the other hand, is restrained to the linear combinations of observations that yield to admissible codewords. The improvement provided by the augmented discriminating power can be observed in Figs. 1 and 2 where, for different sizes of the finite field and different number of sources, we report the *failure rate* of the technique *vs.* the number of samples of the observations in log-scale. The failure rate is simply 1 minus the success rate, where the success rate is the number of correctly identified sources divided by the total number of sources. Note that, as mentioned before, a source is considered identified up to a permutation and scaling ambiguity. It is worth noting that, thanks to the properties of the bit-code, even though the scaling ambiguity is still present, it is in practice drastically reduced (see the Appendix for more details). We observe that our technique consistently outperforms the reference technique, thanks to the possibility of eliminating candidate solutions with low entropy on the grounds that they are not codewords. Since the failure rate converges to zero with the number of samples, as we expected, the gain decreases with the length of the sources. However, the introduction of the non-linear code significantly improves the performances for shorter sources, making the separation viable for relatively shorter signals.

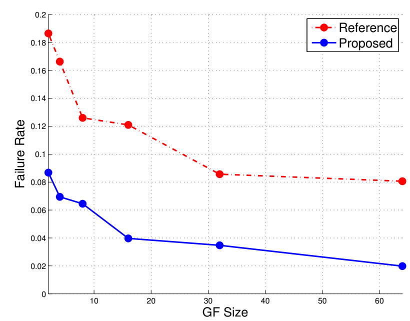

Fig. 3. Comparison between the reference method and the proposed technique for a fixed number of sources and samples. The failure rate is plotted against the size of the finite field.

We also report, in Fig. 3, a comparison of the two techniques with fixed number of sources ($N = 2$) and fixed number of samples ($T = 256$) to observe how the performances of the two methods vary w.r.t. the size of the finite field. We observe that both techniques perform better when they operate within a larger finite field, but the gain of the reference technique stays more or less constant around 6 %.

5 Conclusions and Future Work

In this work we proposed to use a non-linear channel encoding of source signals over a finite field in order to increase the discriminating power of blind source separation methods for linear mixtures in a finite field. In particular, we use an odd-parity bit code, which has the advantage of being very simple to implement. However, these results can be extended to a more general case of a non-linear error detecting code.

The discriminating power is augmented in the sense that the entropy based method will be assisted by the error detecting coding, restraining the estimation of the entropy to the solutions that are admissible in the sense that the reconstructed source is a codeword. This eliminates several solutions that, even if they present low entropy and could be mistakenly identified as sources by the reference technique, cannot be admitted as they are not part of the code. Our experimental results show that the proposed technique consistently outperforms the reference method, especially in the case of sources with a small number of available samples, which is more critical for the entropy-based methods, making the blind source separation more suitable for practical applications, where the number of samples is typically limited by the size of a packet.

These results suggest that a viable future work is to evaluate the performance of the algorithm when the entropy-based method is augmented with a more efficient error detecting code, *i.e.*, a code able to provide better discrimination with lower overhead. This could allow the implementation of a transmission system similar to Network Coding, but with a substantially reduced overhead since the combination coefficient used in the mixing functions do not need to be transmitted.

6 Appendix

In this Appendix we shall evaluate the probability of a random linear combination of N sources encoded with an odd-parity bit code of being a codeword itself. This probability is useful to assess the augmented discriminating power provided by the encoding w.r.t. the separation of the sources.

Let C be the application associating a codeword to each element of GF $\left(2^{b-1}\right)$: C: GF $\left(2^{b-1}\right) \rightarrow$ GF $\left(2^{b}\right)$. This simply amounts to add a odd-parity bit to the binary representation of the element. Let $\mathcal{I}_C \subset$ GF $\left(2^b\right)$ be the image of C. One important property of C is that, by construction, $0 \notin \mathcal{I}_C$ and $1 \in \mathcal{I}_C$, $\forall b \in \mathbb{N}$. Also, it is easy to see that $\|\mathcal{I}_C\| = \frac{\|\mathrm{GF}(2^b)\|}{2} = 2^{b-1}$, *i.e.*, half of the elements of GF $\left(2^b\right)$ are codewords. We can therefore infer that, if a value α is drown from a uniform distribution over GF $\left(2^b\right)$, $\mathrm{P}\left\{\alpha \in \mathcal{I}_C\right\} = \frac{1}{2}$.

Let us consider a monomial $x = \alpha s$, with $\alpha \in$ GF $\left(2^b\right)$ and $s \in \mathcal{I}_C$. In order to evaluate the probability $\mathrm{P}\left\{x \in \mathcal{I}_C\right\}$, we decompose the sample space in the following way:

$$\begin{aligned} \mathrm{P}\left\{x \in \mathcal{I}_C\right\} = \; & \mathrm{P}\left\{\alpha s \in \mathcal{I}_C | \alpha = 0\right\} \mathrm{P}\left\{\alpha = 0\right\} \\ & + \mathrm{P}\left\{\alpha s \in \mathcal{I}_C | \alpha = 1\right\} \mathrm{P}\left\{\alpha = 1\right\} \\ & + \mathrm{P}\left\{\alpha s \in \mathcal{I}_C | \alpha \neq 0,1\right\} \mathrm{P}\left\{\alpha \neq 0,1\right\}. \end{aligned} \quad (1)$$

We operate this decomposition on the base of the properties of elements 0 and 1 w.r.t. multiplication: $0 \cdot s = 0 \notin \mathcal{I}_C$ and $1 \cdot s = s \in \mathcal{I}_C$ with probability 1. In the remaining cases, *i.e.*, when $\alpha \neq 0$ and $\alpha \neq 1$, it is easy to verify that the probability of the monomial being a codeword is $\frac{1}{2}$, based on the fact that the product of a scalar other than 0 for all the other elements of the finite amounts to a reordering of the elements. The probability of $\alpha = 0$ (respectively, $\alpha = 1$) being one out of the number of elements in GF $\left(2^b\right)$, we can rewrite Eq. (1) as: $\mathrm{P}\left\{x \in \mathcal{I}_C\right\} = 0 \cdot \frac{1}{2^b} + 1 \cdot \frac{1}{2^b} + \frac{1}{2} \cdot \frac{2^b - 2}{2^b} = \frac{1}{2}$.

The properties of elements zero and one w.r.t. multiplications become relevant if we consider, instead of the product of two scalars, the product of a scalar by a vector of T elements, *i.e.*, $\mathbf{x} = \alpha \mathbf{s}$ with $\alpha \in$ GF $\left(2^b\right)$ and $\mathbf{s} \in \mathcal{I}_C^T$. We define a *codevector* as being any vector of GF $\left(2^b\right)^T$ such that each one of its elements is a codeword. In this case we observe that $\forall t \in \{1 \ldots T\}$, $1 \cdot s_t \notin \mathcal{I}_C$ and $0 \cdot s_t \in \mathcal{I}_C$. In other words, if $\alpha = 0$ or $\alpha = 1$, the events $\alpha s_t \in \mathcal{I}_C$ for all t are not independent, whereas given any other α, these events are independent with probability $\frac{1}{2}$.

We can therefore operate the same partition as in Eq. (1), and write the probability of $\mathbf{x} \in \mathcal{I}_C^T$ as a function of the finite field size 2^b, or equivalently of b, and the vector length T:

$$\pi_1(b, T) \triangleq \mathrm{P}\left\{\mathbf{x} \in \mathcal{I}_C^T\right\} = 2^{-b} + 2^{-T}(1 - 2^{1-b}). \quad (2)$$

The function $\pi_1(b, T)$ is defined as the probability of a single (vector) monomial $\alpha \mathbf{s}$ of being a codevector. Let us now evaluate the probability $\pi_2(b, T)$ of a mixture of two sources of being a codevector. Note that all sources are by hypothesis codevectors. Let $\mathbf{x_2} = \alpha_1 \mathbf{s_1} + \alpha_2 \mathbf{s_2}$. If we operate a decomposition analogous to that of Eq. (2) we obtain:

$$\begin{aligned} \pi_2(b, T) & \triangleq \mathrm{P}\left\{\mathbf{x_2} \in \mathcal{I}_C^T\right\} \\ & = \; 2^{2(1-b)-1} + 2^{-T}\left(1 - 2^{2(1-b)}\right). \end{aligned} \quad (3)$$

For the case of a linear combination of N sources,let us consider a vector $\mathbf{x_N} = \sum_{n=1}^{N} \alpha_n \mathbf{s}_n$; the expression in Eq. (3) can be generalized for N sources as follows:

$$\begin{aligned} \pi_N(b, T) & \triangleq \mathrm{P}\left\{\mathbf{x_N} \in \mathcal{I}_C^T\right\} \\ & = 2^{N(1-b)-1} + 2^{-T}\left(1 - 2^{N(1-b)}\right). \end{aligned}$$

This probability converges to $\pi_N(b) = 2^{N(1-b)-1}$ for $T \rightarrow \infty$, therefore we observe that the probability of a random combination of N encoded sources in GF $\left(2^{b-1}\right)$ decreases with the size of the finite field.

This probability can be interpreted as follows: if the algorithm were based exclusively on the bit code, *i.e.*, if it were to identify as sources any codewords it finds, for sufficiently long sources it would have a rate of false-positive equal of $\pi_N(b)$. Of course our method does not rely solely on the discriminating power of the code, but the coding is used to drastically reduce the search space of the entropy-based method, reducing the running time and improving the success rate.

References

[1] P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent Component Analysis and Applications*, 1st ed. Academic Press, 2010.

[2] M. Castella, J.-C. Pesquet, and A. Petropulu, “A family of frequency- and time-domain contrasts for blind separation of convolutive mixtures of temporally dependent signals,” *IEEE Transactions on Signal Processing*, vol. 53, no. 1, pp. 107–120, Jan. 2005.

[3] P. Comon, “Independent component analysis, a new concept?” *Signal Processing (Elsevier Science)*, vol. 36, no. 3, pp. 287–314, Apr. 1994. [Online]. Available: http://dx.doi.org/10.1016/0165-1684(94)90029-9

[4] A. Hyvärinen and E. Oja, “Independent component analysis: algorithms and applications,” *Elsevier Journal on Neural Networks*, vol. 13, no. 4-5, pp. 411–430, May 2000. [Online]. Available: http://dx.doi.org/10.1016/S0893-6080(00)00026-5

[5] A. Yeredor, “ICA in Boolean XOR mixtures,” in *Proceedings of Springer-Verlag Internantional Conference on Independent Component Analysis*, 2007, pp. 827–835.

[6] ——, “Independent component analysis over Galois Fields of prime order,” *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 5342–5359, Aug. 2011.

[7] R. Ahlswede, N. Cai, S.-Y. Li, and R. Yeung, “Network information flow,” *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000.

[8] S.-Y. R. Li, R. W. Yeung, and N. Cai, “Linear network coding,” *IEEE Transactions on Information Theory*, vol. 49, no. 2, pp. 371–381, Feb. 2003.

[9] R. Koetter and M. Médard, “An algebraic approach to network coding,” *IEEE/ACM Transactions on Networking*, vol. 11, no. 5, pp. 782–795, Oct. 2003.

[10] T. Ho, M. Médard, J. Shi, M. Effros, and D. Karger, “On randomized network coding,” in *Proceedings of IEEE International Symposium on Information Theory*, Kanagawa, Japan, Jun. 2003.

[11] P. Chou, Y. Wu, and K. Jain, “Practical network coding,” in *Proceedings of Allerton Conference on Communication Control and Computing*, Monticello, IL, USA, Oct. 2003.

[12] D. Vukobratović and V. Stanković, “Unequal error protection random linear coding for multimedia communications,” in *Proceedings of IEEE Workshop on Multimedia Signal Processing*, Saint-Malo, France, Oct. 2010.

[13] N. Thomos, J. Chakareski, and P. Frossard, “Prioritized distributed video delivery with randomized network coding,” *IEEE Transactions on Multimedia*, vol. 13, no. 4, pp. 776–787, Aug. 2011.

[14] I. D. Nemoianu, C. Greco, M. Cagnazzo, and B. Pesquet-Popescu, “A framework for joint multiple description coding and network coding over wireless ad-hoc networks,” in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, Kyoto, Japan, Mar. 2012.

[15] C. Greco, I. D. Nemoianu, M. Cagnazzo, and B. Pesquet-Popescu, “A network coding scheduling for multiple description video streaming over wireless networks,” in *Proceedings of European Signal Processing Conference*, Bucharest, Romania, Aug. 2012.

[16] D. Vukobratovic and V. Stankovic, “Unequal error protection random linear coding strategies for erasure channels,” *IEEE Transactions on Communications*, vol. 60, no. 5, pp. 1243 – 1252, may 2012.

[17] E. Magli, M. Wang, P. Frossard, and A. Markopoulou, “Network Coding meets multimedia: A review,” *IEEE Transactions on Multimedia*, Oct. 2012, accepted for publication.

[18] H. Gutch, P. Gruber, and F. Theis, “ICA over finite fields,” in *Proceedings of Springer-Verlag International Conference on Latent Variable Analysis and Signal Separation*, 2010, pp. 645–652.

[19] H. W. Gutch, P. Gruber, A. Yeredor, and F. J. Theis, “ICA over finite field–Separability and algorithms,” *Signal Processing (Elsevier Science)*, vol. 92, no. 8, pp. 1796–1808, 2012.