



**ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ**

**ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ  
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ**

**ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ**

**Ανάπτυξη και αξιολόγηση αλγορίθμων συμπερασμού κατά  
Bayes κατάλληλων για φασματικό διαχωρισμό**

**Βασίλειος Χ. Αγγελόπουλος  
Πάυλος Π. Καφούρης**

**Επιβλέποντες: Σέργιος Θεοδωρίδης, Καθηγητής ΕΚΠΑ  
Κωνσταντίνος Κουτρούμπας, Κύριος Ερευνητής ΕΑΑ  
Αθανάσιος Ροντογιάννης, Κύριος Ερευνητής ΕΑΑ**

**ΑΘΗΝΑ**

**ΑΥΓΟΥΣΤΟΣ 2012**

## **ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ**

Ανάπτυξη και αξιολόγηση αλγορίθμων συμπερασμού κατά Bayes κατάλληλων για φασματικό διαχωρισμό

**Βασίλειος Χ. Αγγελόπουλος**

**A.M.: 1115 2005 00015**

**Πάυλος Π. Καφούρης**

**A.M.: 1115 2007 00241**

**ΕΠΙΒΛΕΠΟΝΤΕΣ:** Σέργιος Θεοδωρίδης, Καθηγητής ΕΚΠΑ  
Κωνσταντίνος Κουτρούμπας, Κύριος Ερευνητής ΕΑΑ  
Αθανάσιος Ροντογιάννης, Κύριος Ερευνητής ΕΑΑ

## ΠΕΡΙΛΗΨΗ

Η παρούσα εργασία, επικεντρώνεται στο πρόβλημα του ημί-επιβλεπόμενου υπέρ-φασματικού διαχωρισμού, όπου είναι γνωστές οι φασματικές υπογραφές κάποιων υλικών και ο στόχος κατά την ανάλυση (της φασματικής υπογραφής) ενός συγκεκριμένου εικονοστοιχείου είναι ο προσδιορισμός τόσο των υλικών που (μέσω των φασματικών υπογραφών τους) συνθέτουν το εικονοστοιχείο, όσο και του διανύσματος που περιέχει τα ποσοστά με τα οποία συμμετέχουν τα παραπάνω υλικά στη σύνθεση αυτή. Το γραμμικό μοντέλο μίξης είναι αυτό που υιοθετείται για κάθε εικονοστοιχείο της υπό εξέταση υπερφασματικής εικόνας. Με βάση το μοντέλο αυτό, αναπτύσσεται μία ιεραρχική προσέγγιση κατά Bayes κατάλληλη για ημί-επιβλεπόμενο υπέρ-φασματικό διαχωρισμό, όπου έχουν ανατεθεί κατάλληλες εκ των προτέρων κατανομές πιθανότητας στις εμπλεκόμενες υπό εκτίμηση παραμέτρους, οι οποίες μοντελοποιούν την ιδιότητα της αραιότητας (που αντικατοπτρίζει το γεγονός ότι, στην πράξη, μόνο λίγα από τα διαθέσιμα υλικά συμμετέχουν στο σχηματισμό ενός εικονοστοιχείου) και της μη αρνητικότητας του διανύσματος των ποσοστών. Στη συνέχεια, εξάγεται ένας νέος επαναληπτικός αλγόριθμος συμπερασμού κατά Bayes, με όνομα BI-ICE-single, ο οποίος παράγει αραιές εκτιμήσεις του διανύσματος ποσοστών, ικανοποιώντας τον περιορισμό της μη αρνητικότητας. Τέλος, περιγράφεται μια νέα απλή τεχνική, η οποία λαμβάνει υπ' όψιν της την πιθανή χωρική συσχέτιση μεταξύ γειτονικών εικονοστοιχείων της υπερφασματικής εικόνας, κατά την εφαρμογή του φασματικού διαχωρισμού (είτε μέσω του BI-ICE-single, είτε μέσω άλλου αλγορίθμου του ίδιου σκοπού, π.χ. τον BI-ICE [1]). Τα πειραματικά αποτελέσματα δείχνουν ότι ο BI-ICE-single δεν παρουσιάζει την ακρίβεια στην εκτίμηση της μεθόδου BI-ICE, στην περίπτωση όπου οι φασματικές υπογραφές των υλικών παρουσιάζουν μεγάλο βαθμό συσχέτισης. Αντίθετα, όταν ο βαθμός συσχέτισης είναι μικρός, οι δύο αλγόριθμοι παρουσιάζουν παρόμοια συμπεριφορά. Επιπλέον, η προτεινόμενη τεχνική αξιοποίησης της συσχέτισης μεταξύ γειτονικών εικονοστοιχείων οδηγεί σε σημαντική εξοικονόμηση υπολογιστικής ισχύος, χωρίς να υστερεί στην ποιότητα σε σχέση με την περίπτωση που η πιθανή χωρική συσχέτιση αγνοείται.

**ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ:** Επεξεργασία υπερφασματικών δεδομένων

**ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ:** υπερφασματική απεικόνιση, ιεραρχικό Bayesian μοντέλο, αραιός ημί-επιβλεπόμενος φασματικός διαχωρισμός, συμπερασμός κατά Bayes, φασματική υπογραφή

## **ABSTRACT**

In this thesis, the problem of semisupervised hyperspectral unmixing is considered, where the spectral signatures of some materials are known and the aim during the analysis (of the spectral signature) of a specific pixel is to determine both the materials that (through their spectral signatures) contribute to the composition of the pixel and the vector containing the abundance fractions of each one of these materials in the composition. The linear mixture model is the one that is adopted for every pixel of the examined hyperspectral image. Based on this model, a hierarchical Bayesian approach suitable for semisupervised hyperspectral unmixing is proposed, where suitable priors are selected for the model parameters such that, on the one hand, they favor sparse solutions for the abundance vector (which indicate the fact that, in practice, only few of the materials are presented in a specific pixel), while on the other hand they ensure the non-negativity of the abundances. Then, a new Bayesian inference iterative scheme, named BI-ICE-single, is developed, which produces sparse estimations for the abundance vector that satisfy the non-negativity constraint. Finally, a new simple technique is described, which takes into account the possible spatial correlation between adjacent pixels of hyperspectral image during the application of spectral unmixing (using BI-ICE-single, or another algorithm of the same purpose, such as BI-ICE [1]).

Experimental results illustrate that the BI-ICE-single algorithm does not present the estimation accuracy of BI-ICE method, in the case where the spectral signatures of the endmembers are highly correlated. On the contrary, the two algorithms exhibit similar performance, when the spectral signatures of the endmembers are not highly correlated. In addition, the proposed technique that takes into account the possible spatial correlation between adjacent pixels offers significant computational savings, without leading to inferior quality results compared to the case where the possible spatial correlation is ignored.

**SUBJECT AREA:** Hyperspectral data processing

**KEYWORDS:** hyperspectral imagery, hierarchical Bayesian model, sparse semisupervised spectral unmixing, Bayesian inference, spectral signature

*ΣΤΙΣ ΟΙΚΟΓΕΝΕΙΕΣ ΜΑΣ*

## **ΕΥΧΑΡΙΣΤΙΕΣ**

Για τη διεκπεραίωση της παρούσας Πτυχιακής Εργασίας, θα θέλαμε να ευχαριστήσουμε τον επιβλέποντα καθηγητή, κ. Σέργιο Θεοδωρίδη, καθώς και τους συνεργάτες του, κύριους ερευνητές του ΕΑΑ κ. Κωνσταντίνο Κουτρούμπα, κ. Αθανάσιο Ροντογιάννη και διδάκτορα κ. Κωνσταντίνο Θέμελη, για τη συνεργασία και την πολύτιμη συμβολή τους, τόσο στην επιλογή του θέματος της εργασίας, όσο και στην επιτυχή ολοκλήρωσή της.

# ΠΕΡΙΕΧΟΜΕΝΑ

<b>ΠΡΟΛΟΓΟΣ</b> .....	<b>14</b>
<b>1. ΕΙΣΑΓΩΓΗ</b> .....	<b>15</b>
1.1 Γενικά για δεδομένα απεικονιστικής φασματοσκοπίας .....	15
1.2 Φασματικός διαχωρισμός .....	17
1.2.1 Επιβλεπόμενος, ημί-επιβλεπόμενος και μη επιβλεπόμενος φασματικός διαχωρισμός .....	18
1.2.2 Τα βήματα της διαδικασίας του υπέρ-φασματικού διαχωρισμού.....	18
1.2.3 Φυσικοί περιορισμοί .....	18
1.2.4 Προτεινόμενη προσέγγιση κατά Bayes.....	19
1.3 Συμβολισμοί .....	22
<b>2. ΤΟ ΠΡΟΒΛΗΜΑ ΤΟΥ ΑΡΑΙΟΥ ΗΜΙ-ΕΠΙΒΛΕΠΟΜΕΝΟΥ ΦΑΣΜΑΤΙΚΟΥ ΔΙΑΧΩΡΙΣΜΟΥ</b> .....	<b>23</b>
2.1 Ορισμοί.....	23
2.2 Διατύπωση του προβλήματος .....	23
<b>3. ΙΕΡΑΡΧΙΚΟ ΜΟΝΤΕΛΟ ΚΑΤΑ ΒΑΥΕΣ</b> .....	<b>26</b>
3.1 Εισαγωγή .....	26
3.2 Ανάλυση του ιεραρχικού μοντέλου .....	27
3.2.1 Συνάρτηση πιθανοφάνειας.....	27
3.2.2 Εκ των προτέρων κατανομές των παραμέτρων.....	27
3.2.3 Εκ των προτέρων κατανομές των υπέρ-παραμέτρων .....	30
<b>4. Ο ΠΡΟΤΕΙΝΟΜΕΝΟΣ ΑΛΓΟΡΙΘΜΟΣ ΣΥΜΠΕΡΑΣΜΟΥ ΚΑΤΑ ΒΑΥΕΣ</b> .....	<b>34</b>
4.1 Εισαγωγή .....	34
4.2 Υπό συνθήκη εκ των υστέρων κατανομές.....	34
4.3 Ο αλγόριθμος BI-ICE-single .....	37
4.3.1 Υπό συνθήκη μέση τιμή της $p(\vec{w}   \vec{y}, \vec{\gamma}, \lambda, \beta)$ .....	38
4.3.2 Υπό συνθήκη μέση τιμή της $p(\beta   \vec{y}, \vec{w}, \vec{\gamma}, \lambda)$ .....	39

4.3.3	Υπό συνθήκη μέση τιμή της $p(\gamma_i   \bar{y}, w_i, \lambda, \beta)$ .....	39
4.3.4	Υπό συνθήκη μέση τιμή της $p(\lambda   \bar{y}, w, \gamma, \beta)$ .....	40
4.4	Παρουσίαση του BI-ICE-single αλγορίθμου .....	40
4.5	Ενσωμάτωση προσθετικού περιορισμού στο γραμμικό μοντέλο μίξης .....	42
<b>5.</b>	<b>ΦΑΣΜΑΤΙΚΟΣ ΔΙΑΧΩΡΙΣΜΟΣ ΜΕ ΑΞΙΟΠΟΙΗΣΗ ΤΗΣ ΧΩΡΙΚΗΣ ΣΥΣΧΕΤΙΣΗΣ ΜΕΤΑΞΥ ΓΕΙΤΟΝΙΚΩΝ ΕΙΚΟΝΟΣΤΟΙΧΕΙΩΝ .....</b>	<b>44</b>
<b>6.</b>	<b>ΑΠΟΤΕΛΕΣΜΑΤΑ ΠΕΙΡΑΜΑΤΙΚΗΣ ΜΕΛΕΤΗΣ .....</b>	<b>47</b>
6.1	Αποτελέσματα προσομοίωσης σε συνθετικά δεδομένα .....	47
6.2	Αποτελέσματα προσομοίωσης σε πραγματικά δεδομένα .....	57
<b>7.</b>	<b>ΣΥΜΠΕΡΑΣΜΑΤΑ .....</b>	<b>65</b>
	<b>ΠΙΝΑΚΑΣ ΟΡΟΛΟΓΙΑΣ .....</b>	<b>66</b>
	<b>ΣΥΝΤΜΗΣΕΙΣ – ΑΡΚΤΙΚΟΛΕΞΑ – ΑΚΡΩΝΥΜΙΑ .....</b>	<b>68</b>
	<b>ΠΑΡΑΡΤΗΜΑ – ΚΩΔΙΚΑΣ MATLAB .....</b>	<b>69</b>
	<b>ΑΝΑΦΟΡΕΣ .....</b>	<b>92</b>



## ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ

Σχήμα 1: Κατευθυνόμενο μη κυκλικό γράφημα του προτεινόμενου Bayesian μοντέλου. Οι παράμετροι που δεν είναι τυχαίες μεταβλητές εμφανίζονται σε ορθογώνια πλαίσια .32

Σχήμα 2: Εκτίμηση των στοιχείων του διανύσματος ποσοστών  $\bar{w}$  , καθώς εξελίσσεται η εκτέλεση των αλγορίθμων (α) BI-ICE και (β) BI-ICE-single. Οι αλγόριθμοι εφαρμόζονται σε προσομοιωμένα δεδομένα, που έχουν δημιουργηθεί χρησιμοποιώντας ένα μητρώο φασματικών υπογραφών, οι οποίες παρουσιάζουν υψηλή συσχέτιση μεταξύ τους. Έχει προστεθεί λευκός θόρυβος (SNR = 25 dB). Διακεκομμένες γραμμές: Πραγματικές τιμές μη μηδενικών συνιστωσών. Συμπαγείς γραμμές: Εκτιμώμενες τιμές .....48

Σχήμα 3: Εκτίμηση των στοιχείων του διανύσματος ποσοστών  $\bar{w}$  , καθώς εξελίσσεται η εκτέλεση των αλγορίθμων (α) BI-ICE και (β) BI-ICE-single. Οι αλγόριθμοι εφαρμόζονται σε προσομοιωμένα δεδομένα, που έχουν δημιουργηθεί χρησιμοποιώντας ένα μητρώο φασματικών υπογραφών, οι οποίες ακολουθούν την κανονική κατανομή. Έχει προστεθεί λευκός θόρυβος (SNR = 25 dB). Διακεκομμένες γραμμές: Πραγματικές τιμές μη μηδενικών συνιστωσών. Συμπαγείς γραμμές: Εκτιμώμενες τιμές .....49

Σχήμα 4: Εκτίμηση των στοιχείων του διανύσματος ποσοστών  $\bar{w}$  , καθώς εξελίσσεται η εκτέλεση των αλγορίθμων (α) BI-ICE και (β) BI-ICE-single. Οι αλγόριθμοι εφαρμόζονται σε προσομοιωμένα δεδομένα, που έχουν δημιουργηθεί χρησιμοποιώντας ένα μητρώο φασματικών υπογραφών, οι οποίες είναι ανεξάρτητα και ομοιόμορφα κατανεμημένες. Έχει προστεθεί λευκός θόρυβος (SNR = 25 dB). Διακεκομμένες γραμμές: Πραγματικές τιμές μη μηδενικών συνιστωσών. Συμπαγείς γραμμές: Εκτιμώμενες τιμές .....49

Σχήμα 5: Μέσο τετραγωνικό σφάλμα που παράγεται από τους BI-ICE και BI-ICE-single, για διάφορους βαθμούς αραιότητας. Οι δύο αλγόριθμοι εφαρμόζονται σε προσομοιωμένα δεδομένα που προκύπτουν από το μητρώο  $\Phi$ , ενώ έχει προστεθεί λευκός θόρυβος (SNR = 20 dB).....50

Σχήμα 6: Μέσο τετραγωνικό σφάλμα που παράγεται από τους BI-ICE και BI-ICE-single, για διάφορους βαθμούς αραιότητας. Οι δύο αλγόριθμοι εφαρμόζονται σε προσομοιωμένα δεδομένα που προκύπτουν από το μητρώο  $\Phi_1$ , ενώ έχει προστεθεί λευκός θόρυβος (SNR = 20 dB).....51

Σχήμα 7: Μέσο τετραγωνικό σφάλμα που παράγεται από τους BI-ICE και BI-ICE-single, για διάφορους βαθμούς αραιότητας. Οι δύο αλγόριθμοι εφαρμόζονται σε

προσομοιωμένα δεδομένα που προκύπτουν από το μητρώο $\Phi_2$ , ενώ έχει προστεθεί λευκός θόρυβος (SNR = 20 dB).....	51
Σχήμα 8: Μέσο τετραγωνικό σφάλμα που προκύπτει λαμβάνοντας υπ' όψιν μόνο τα υλικά που συνθέτουν το εικονοστοιχείο, το οποίο παράγεται από τους BI-ICE και BI-ICE-single, για διαφόρους βαθμούς αραιότητας. Οι δύο αλγόριθμοι εφαρμόζονται σε προσομοιωμένα δεδομένα που προκύπτουν από το μητρώο $\Phi$ , ενώ έχει προστεθεί λευκός θόρυβος (SNR = 20 dB).....	52
Σχήμα 9: Μέσο τετραγωνικό σφάλμα που προκύπτει λαμβάνοντας υπ' όψιν μόνο τα υλικά που συνθέτουν το εικονοστοιχείο, το οποίο παράγεται από τους BI-ICE και BI-ICE-single, για διαφόρους βαθμούς αραιότητας. Οι δύο αλγόριθμοι εφαρμόζονται σε προσομοιωμένα δεδομένα που προκύπτουν από το μητρώο $\Phi_1$ , ενώ έχει προστεθεί λευκός θόρυβος (SNR = 20 dB).....	53
Σχήμα 10: Μέσο τετραγωνικό σφάλμα που προκύπτει λαμβάνοντας υπ' όψιν μόνο τα υλικά που συνθέτουν το εικονοστοιχείο, το οποίο παράγεται από τους BI-ICE και BI-ICE-single, για διαφόρους βαθμούς αραιότητας. Οι δύο αλγόριθμοι εφαρμόζονται σε προσομοιωμένα δεδομένα που προκύπτουν από το μητρώο $\Phi_2$ , ενώ έχει προστεθεί λευκός θόρυβος (SNR = 20 dB).....	53
Σχήμα 11: Μέσο τετραγωνικό σφάλμα που παράγεται από τους BI-ICE και BI-ICE-single σαν συνάρτηση του λόγου σήματος-προς-θόρυβο (dB), όταν (οι αλγόριθμοι) αυτοί εφαρμόζονται σε προσομοιωμένα δεδομένα που προκύπτουν από το μητρώο $\Phi$ , ενώ έχει προστεθεί λευκός θόρυβος. Ο βαθμός αραιότητας $\xi$ ισούται με 8.....	54
Σχήμα 12: Μέσο τετραγωνικό σφάλμα που παράγεται από τους BI-ICE και BI-ICE-single σαν συνάρτηση του λόγου σήματος-προς-θόρυβο (dB), όταν (οι αλγόριθμοι) αυτοί εφαρμόζονται σε προσομοιωμένα δεδομένα που προκύπτουν από το μητρώο $\Phi_1$ , ενώ έχει προστεθεί λευκός θόρυβος. Ο βαθμός αραιότητας $\xi$ ισούται με 8. ....	54
Σχήμα 13: Μέσο τετραγωνικό σφάλμα που παράγεται από τους BI-ICE και BI-ICE-single σαν συνάρτηση του λόγου σήματος-προς-θόρυβο (dB), όταν (οι αλγόριθμοι) αυτοί εφαρμόζονται σε προσομοιωμένα δεδομένα που προκύπτουν από το μητρώο $\Phi_2$ , ενώ έχει προστεθεί λευκός θόρυβος. Ο βαθμός αραιότητας $\xi$ ισούται με 8. ....	55
Σχήμα 14: Μέσο τετραγωνικό σφάλμα που παράγεται από τους BI-ICE και BI-ICE-single σαν συνάρτηση του λόγου σήματος-προς-θόρυβο (dB), όταν (οι αλγόριθμοι) αυτοί εφαρμόζονται σε προσομοιωμένα δεδομένα που προκύπτουν από το μητρώο $\Phi$ ,	

ενώ έχει προστεθεί λευκός θόρυβος και λαμβάνονται υπ' όψιν μόνο τα υλικά που υπάρχουν στο εικονοστοιχείο. Ο βαθμός αραιότητας  $\xi$  ισούται με 8.....56

Σχήμα 15: Μέσο τετραγωνικό σφάλμα που παράγεται από τους BI-ICE και BI-ICE-single σαν συνάρτηση του λόγου σήματος-προς-θόρυβο (dB), όταν (οι αλγόριθμοι) αυτοί εφαρμόζονται σε προσομοιωμένα δεδομένα που προκύπτουν από το μητρώο  $\Phi_1$ , ενώ έχει προστεθεί λευκός θόρυβος και λαμβάνονται υπ' όψιν μόνο τα υλικά που υπάρχουν στο εικονοστοιχείο. Ο βαθμός αραιότητας  $\xi$  ισούται με 8.....56

Σχήμα 16: Μέσο τετραγωνικό σφάλμα που παράγεται από τους BI-ICE και BI-ICE-single σαν συνάρτηση του λόγου σήματος-προς-θόρυβο (dB), όταν (οι αλγόριθμοι) αυτοί εφαρμόζονται σε προσομοιωμένα δεδομένα που προκύπτουν από το μητρώο  $\Phi_2$ , ενώ έχει προστεθεί λευκός θόρυβος και λαμβάνονται υπ' όψιν μόνο τα υλικά που υπάρχουν στο εικονοστοιχείο. Ο βαθμός αραιότητας  $\xi$  ισούται με 8.....57

Σχήμα 17: Γραφική απεικόνιση του συνολικού αριθμού επαναλήψεων κατά τη διάρκεια του πειράματος, για τον αλγόριθμο BI-ICE με και χωρίς χωρική συσχέτιση (μπλε συνεχής γραμμή και κόκκινη διακεκομμένη γραμμή, αντίστοιχα).....62

Σχήμα 18: Γραφική απεικόνιση του συνολικού αριθμού επαναλήψεων κατά τη διάρκεια του πειράματος, για τον αλγόριθμο BI-ICE-single με και χωρίς χωρική συσχέτιση (μπλε συνεχής γραμμή και κόκκινη διακεκομμένη γραμμή, αντίστοιχα).....62

## ΚΑΤΑΛΟΓΟΣ ΕΙΚΟΝΩΝ

Εικόνα 1: Βασικές έννοιες της απεικονιστικής φασματοσκοπίας, [3].....	16
Εικόνα 2: Υπό-εικόνα της 150 <sup>ης</sup> ζώνης συχνοτήτων του υπερφασματικού συνόλου δεδομένων Cuprite AVIRIS.....	58
Εικόνα 3: Εκτιμώμενες τιμές ποσοστών για τέσσερα υλικά, χρησιμοποιώντας τον αλγόριθμο BI-ICE χωρίς χωρική συσχέτιση.....	59
Εικόνα 4: Εκτιμώμενες τιμές ποσοστών για τέσσερα υλικά, χρησιμοποιώντας τον αλγόριθμο BI-ICE με χωρική συσχέτιση .....	60
Εικόνα 5: Εκτιμώμενες τιμές ποσοστών για τέσσερα υλικά, χρησιμοποιώντας τον αλγόριθμο BI-ICE-single χωρίς χωρική συσχέτιση .....	60
Εικόνα 6: Εκτιμώμενες τιμές ποσοστών για τέσσερα υλικά, χρησιμοποιώντας τον αλγόριθμο BI-ICE-single με χωρική συσχέτιση.....	61

## ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

Πίνακας 1: Ο αλγόριθμος BI-ICE-single.....	41
Πίνακας 2: Ο αλγόριθμος BI-ICE-single με αξιοποίηση της πληροφορίας της πιθανής χωρικής συσχέτισης .....	45
Πίνακας 3: Μέσος αριθμός επαναλήψεων, συνολικός χρόνος εκτέλεσης και μέσος χρόνος εκτέλεσης ανά εικονοστοιχείο, για τους αλγορίθμους BI-ICE και BI-ICE-single, με και χωρίς χωρική συσχέτιση .....	63

## **ΠΡΟΛΟΓΟΣ**

Η παρούσα Πτυχιακή Εργασία εκπονήθηκε στην Αθήνα κατά το ακαδημαϊκό έτος 2012. Αποτελεί αναπόσπαστο κομμάτι για την απόκτηση του πτυχίου μας και διεξήχθη κατά το τελευταίο έτος φοίτησής μας ως προπτυχιακοί φοιτητές στο τμήμα Πληροφορικής και Τηλεπικοινωνιών του Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών. Στην επιτυχή περάτωση της εργασίας αυτής, ήταν σημαντική η βοήθεια και υποστήριξη του επιβλέποντος καθηγητή μας, κ. Θεοδωρίδη Σέργιου και των συνεργατών του, κ. Κουτρούμπα Κωνσταντίνου, κ. Ροντογιάννη Αθανάσιου και κ. Θέμελη Κωνσταντίνου.

## 1. ΕΙΣΑΓΩΓΗ

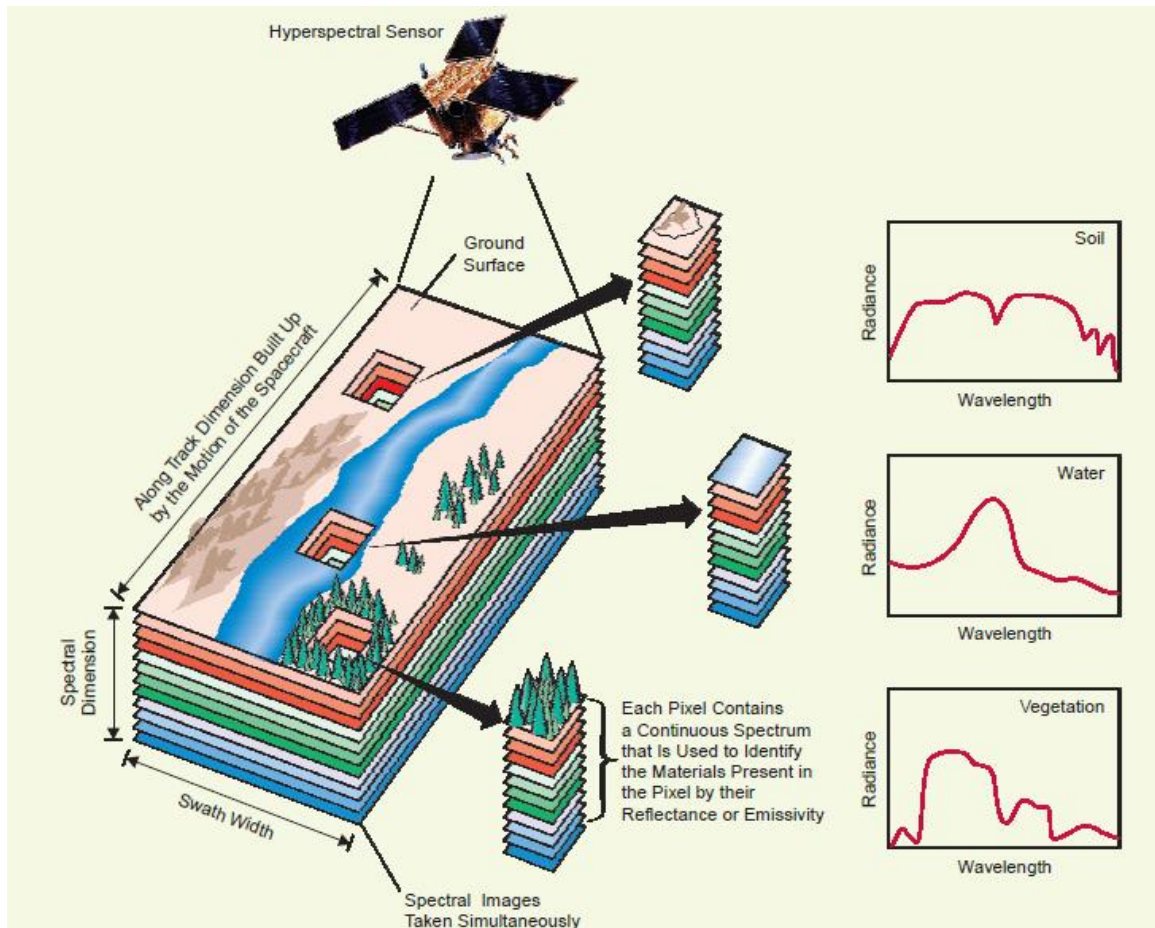
### 1.1 Γενικά για δεδομένα απεικονιστικής φασματοσκοπίας

Τα τελευταία χρόνια, η υπερφασματική απεικόνιση (hyperspectral imagery) χρησιμοποιείται ευρέως σε πολλά πεδία εφαρμογών που παρουσιάζουν αυξημένο ερευνητικό ενδιαφέρον, όπως είναι η παρακολούθηση αλλαγών που παρατηρούνται στο περιβάλλον και η ταξινόμηση του τύπου εδάφους (terrain classifications) [2]-[4]. Ο όρος «επεξεργασία υπερφασματικής εικόνας» περιλαμβάνει πολλά επιμέρους τεχνικά ζητήματα, όπως είναι η ταξινόμηση και η κατάτμηση της εικόνας, η ανίχνευση-εντοπισμός συγκεκριμένων στόχων μέσα σ' αυτήν, καθώς και η διαδικασία φασματικού διαχωρισμού (spectral unmixing - SU) της εικόνας, με την οποία θα ασχοληθούμε εκτενέστερα στη συνέχεια.

Στα πλαίσια της ηλεκτρο-οπτικής τηλεπισκόπησης (electro-optical remote sensing) λαμβάνονται πληροφορίες σχετικά με ένα αντικείμενο ή μία σκηνή (scene), χωρίς να υπάρχει φυσική επαφή με αυτά, μέσω κατάλληλων αισθητήρων. Αυτό καθίσταται δυνατό εξαιτίας της ιδιότητας που παρουσιάζουν τα υλικά, που συνθέτουν τα διάφορα αντικείμενα μίας σκηνής, να αντανακλούν, απορροφούν και εκπέμπουν ηλεκτρομαγνητική ακτινοβολία, τόσο σύμφωνα με την μοριακή σύνθεσή τους, όσο και σύμφωνα με το σχήμα τους. Η ακτινοβολία που επιστρέφει στον αισθητήρα καθορίζει την φασματική υπογραφή (spectral signature), ή απλά φάσμα (spectral), η οποία μπορεί να χρησιμοποιηθεί για να χαρακτηρίσει και εντοπίσει, μοναδικά, ένα δεδομένο υλικό, [3].

Η φασματοσκοπία (spectroscopy) ασχολείται με τη μέτρηση, την ανάλυση και την ερμηνεία αυτών των φασμάτων. Ο συνδυασμός της φασματοσκοπίας με τις μεθόδους που χρησιμοποιούνται για την απόκτηση φασματικών πληροφοριών από μεγάλες περιοχές, είναι γνωστός ως απεικονιστική φασματοσκοπία (imaging spectroscopy). Οι βασικές έννοιες που εμπλέκονται στην απεικονιστική φασματοσκοπία, μπορούν να εξηγηθούν ευκολότερα μέσω του παραδείγματος της Εικόνας 1, [3]. Σε αυτήν, παρατηρούμε τις φασματικές εικόνες που λαμβάνονται ταυτόχρονα από έναν φασματικό αισθητήρα, οι οποίες απεικονίζουν μία σκηνή, όπου το έδαφος, το νερό και η βλάστηση είναι τα κύρια υλικά των αντικειμένων της. Αυτά αντανακλούν, απορροφούν και εκπέμπουν ηλεκτρομαγνητική ακτινοβολία. Η ακτινοβολία, που επιστρέφει στον αισθητήρα από ένα συγκεκριμένο υλικό, καθορίζει την φασματική υπογραφή του, η οποία χρησιμοποιείται στη συνέχεια για να εντοπίσει το υλικό αυτό σε άλλες θέσεις (εικονοστοιχεία – pixels) της εικόνας. Επίσης, απεικονίζονται τρία εικονοστοιχεία από διαφορετικές περιοχές της εικόνας, τα οποία αναπαριστώνται από διανύσματα μεγάλης

(φασματικής) διάστασης (spectral dimension) και χρησιμοποιούνται στη συνέχεια για να προσδιορίσουν τα υλικά που απεικονίζονται στο εικονοστοιχείο. Τέλος, τα τρία διαγράμματα παρουσιάζουν την αναλογία της ακτινοβολίας, που εκπέμπει το κάθε ένα υλικό (έδαφος, νερό, βλάστηση), σε διαδοχικά μήκη κύματος (wavelengths) (φασματική υπογραφή).



Εικόνα 1: Βασικές έννοιες της απεικονιστικής φασματοσκοπίας, [3]

Οι υπερφασματικοί αισθητήρες απεικόνισης (hyperspectral imaging sensors) αποτελούν ειδική κατηγορία των αισθητήρων απεικονιστικής φασματοσκοπίας. Στους υπερφασματικούς αισθητήρες, η ζώνη συχνοτήτων (waveband) διαιρείται σε εκατοντάδες συνεχόμενες στενές ζώνες, ιδιότητα που καθιστά τους αισθητήρες αυτούς ικανούς να δειγματοληπτίσουν σε ένα μεγάλο αριθμό συνεχόμενων στενών φασματικών ζωνών. Έτσι, κάθε εικονοστοιχείο (pixel) μιας υπερφασματικής εικόνας αναπαρίσταται από ένα διάνυσμα μεγάλης διάστασης, το οποίο μπορεί να ιδωθεί ως ένα σχεδόν συνεχές φάσμα ακτινοβολίας (δες ξανά την Εικόνα 1). Αυτή η φασματική πληροφορία, αξιοποιούμενη κατάλληλα, μπορεί να καθορίσει διάφορα ενδιαφέροντα χαρακτηριστικά μιας απομακρυσμένης σκηνής (remote scene), [5].



Στις υπερφασματικές εικόνες, είναι αρκετά συνηθισμένο τα εικονοστοιχεία να απεικονίζουν περισσότερα από ένα διαφορετικά υλικά. Η εμφάνιση μεικτών εικονοστοιχείων οφείλεται σε δύο σημαντικούς λόγους. Ο πρώτος λόγος σχετίζεται με την χωρική ανάλυση (spatial resolution) του αισθητήρα. Συγκεκριμένα, όταν η χωρική ανάλυσή του είναι αρκετά χαμηλή, διαφορετικά υλικά απεικονίζονται σε ένα και μόνο εικονοστοιχείο, με αποτέλεσμα η φασματική μέτρηση, που θα προκύψει, να είναι μία σύνθεση των επιμέρους φασμάτων. Αυτό συμβαίνει σε περιπτώσεις πλατφόρμων τηλεπισκόπησης που βρίσκονται σε μεγάλο υψόμετρο ή σε ευρείας ζώνης (wide-area) παρακολούθησεις, όπου έχουμε χαμηλή χωρική ανάλυση. Ο δεύτερος λόγος σχετίζεται με το γεγονός ότι διακριτά υλικά μπορεί να συνδυάζονται και να δημιουργούν μία ομοιογενή μίξη (homogeneous mixture), ενδεχόμενο που μπορεί να συμβεί ανεξάρτητα από τη χωρική ανάλυση του αισθητήρα, [4].

Το γεγονός ότι ορισμένα εικονοστοιχεία απεικονίζουν, συχνά, ένα συνδυασμό πολλών διαφορετικών συστατικών, οδηγεί στην ανάγκη του φασματικού διαχωρισμού των μίξεων (συνδυασμών) αυτών. Οι υπερφασματικοί αισθητήρες, με την συλλογή δεδομένων σε εκατοντάδες φασματικές ζώνες, μπορούν, κατ' αρχήν, μετά από κατάλληλη επεξεργασία, να δώσουν αξιόπιστα αποτελέσματα όσον αφορά τον φασματικό διαχωρισμό, [4].

## 1.2 Φασματικός διαχωρισμός

*Φασματικός διαχωρισμός* (hyperspectral unmixing) είναι η διαδικασία με την οποία το μετρούμενο φάσμα ενός μεικτού εικονοστοιχείου αναλύεται στα συστατικά φάσματα των υλικών (endmembers) που το συνθέτουν. Συνήθως, η διαδικασία αυτή παρέχει και τα ποσοστά (abundances) στα οποία παρουσιάζεται κάθε συστατικό υλικό (μέσω της φασματικής υπογραφής του) μέσα στο εικονοστοιχείο. Τα συστατικά υλικά, συνήθως, αντιστοιχούν σε οικεία μακροσκοπικά (macroscopic) υλικά, όπως το νερό, το έδαφος, το μέταλλο και η βλάστηση.

Ο φασματικός διαχωρισμός είναι μία ειδική περίπτωση του γενικευμένου αντιστρόφου προβλήματος, που υπολογίζει τις παραμέτρους που περιγράφουν ένα αντικείμενο, παρατηρώντας την αλληλεπίδραση του αντικειμένου με ένα σήμα, πριν από την άφιξη στον αισθητήρα. Στην περίπτωση της υπερφασματικής τηλεπισκόπησης, το προσπίπτον αυτό σήμα είναι η ηλεκτρομαγνητική ακτινοβολία, [4].

Στην πράξη, χρησιμοποιείται συχνά το γραμμικό μοντέλο μίξης (linear mixture model - LMM) [6], που θεωρεί ότι το φάσμα ενός μεικτού εικονοστοιχείου είναι ένας γραμμικός συνδυασμός των φασμάτων των συστατικών που το συνθέτουν.

### **1.2.1 Επιβλεπόμενος, ημί-επιβλεπόμενος και μη επιβλεπόμενος φασματικός διαχωρισμός**

Οι υπερφασματικές μέθοδοι (διαχωρισμού) κατηγοριοποιούνται σε επιβλεπόμενες (supervised), ημί-επιβλεπόμενες (semi-supervised) και μη επιβλεπόμενες (unsupervised). Στις επιβλεπόμενες μεθόδους, [7]-[9], οι φασματικές υπογραφές των συστατικών θεωρούνται ότι είναι γνωστές εκ των προτέρων, ενώ στις μη επιβλεπόμενες μεθόδους [10], [11], [12], τα φάσματα των συστατικών εκτιμώνται απευθείας από τα δεδομένα. Τέλος, στις ημί-επιβλεπόμενες μεθόδους [13], [14], είναι διαθέσιμη μία φασματική βιβλιοθήκη, η οποία περιέχει τις φασματικές υπογραφές πολλών υλικών. Ο στόχος των ημί-επιβλεπόμενων μεθόδων, στις οποίες ανήκει και το μοντέλο που αναπτύσσεται στην παρούσα εργασία, είναι να καθορίσουν πόσα και ποια υλικά, από τη φασματική βιβλιοθήκη, είναι παρόντα στο υπό εξέταση μεικτό εικονοστοιχείο, καθώς και να εκτιμήσουν τα ποσοστά συμμετοχής των (φασματικών υπογραφών των) υλικών αυτών στη σύνθεση (της φασματικής υπογραφής) του εικονοστοιχείου.

### **1.2.2 Τα βήματα της διαδικασίας του υπέρ-φασματικού διαχωρισμού**

Η διαδικασία του υπέρ-φασματικού διαχωρισμού περιλαμβάνει δύο φάσεις. Κατά την πρώτη φάση, καθορίζονται οι φασματικές υπογραφές των υλικών που περιέχονται στην υπερφασματική εικόνα, οι οποίες απαρτίζουν τη φασματική βιβλιοθήκη. Μερικοί από τους πιο δημοφιλείς αλγορίθμους εξαγωγής υλικών είναι: ο Pixel Purity Index (PPI), [15], ο N-FINDR, [12] και η μέθοδος Vertex Component Analysis (VCA), [10]. Στη δεύτερη φάση, γίνεται η ανάστροφη διαδικασία. Δηλαδή, εκτιμώνται τα ποσοστά, με τα οποία τα υλικά συμμετέχουν στη σύνθεση ενός εικονοστοιχείου της εν λόγω εικόνας.

### **1.2.3 Φυσικοί περιορισμοί**

Λόγω της φύσης του προβλήματος, τα ποσοστά συμμετοχής των υλικών στη δημιουργία ενός εικονοστοιχείου, θα πρέπει να πληρούν δύο περιορισμούς. Ο πρώτος περιορισμός, που είναι γνωστός ως ο περιορισμός της μη αρνητικότητας (nonnegative

constraint), δηλώνει ότι όλα τα ποσοστά των συστατικών για ένα εικονοστοιχείο θα πρέπει να είναι μη αρνητικοί αριθμοί. Ο δεύτερος, που λέγεται και προσθετικός περιορισμός (additivity constraint), δηλώνει ότι όλα τα ποσοστά θα πρέπει να αθροίζονται στη μονάδα. Κάτω από αυτούς τους περιορισμούς, ο φασματικός διαχωρισμός διατυπώνεται ως ένα κυρτό (convex) πρόβλημα βελτιστοποίησης (optimization), που μπορεί να αντιμετωπιστεί χρησιμοποιώντας επαναληπτικές μεθόδους, όπως η μέθοδος των ελαχίστων τετραγώνων, που λαμβάνει υπ' όψιν όλους τους περιορισμούς (fully constrained least squares method), [7], ή αριθμητικές μεθόδους βελτιστοποίησης, π.χ. [16]. Εκτός από τις παραπάνω μεθόδους, έχουν προταθεί και Bayesian μέθοδοι για την αντιμετώπιση του προβλήματος αυτού, όπως π.χ. το σχήμα δειγματοληψίας του Gibbs, που εφαρμόζεται στο ιεραρχικό Bayesian μοντέλο του [13].

Εστιάζοντας, στη συνέχεια στον ημί-επιβλεπόμενο φασματικό διαχωρισμό, ένα επιπλέον χαρακτηριστικό, που απαντάται εδώ, είναι αυτό της **αραιότητας** (sparsity) του διανύσματος των ποσοστών συμμετοχής των υλικών στη δημιουργία ενός εικονοστοιχείου (abundance vector)<sup>1</sup> (δηλ., το εν λόγω διάνυσμα θα έχει πολύ λίγες μη μηδενικές συνιστώσες), απόρροια του γεγονότος ότι σε ένα εικονοστοιχείο απεικονίζεται συνήθως ένας μικρός αριθμός υλικών, σε σχέση με το συνολικό αριθμό των υλικών που απεικονίζονται σε όλα τα εικονοστοιχεία της εικόνας. Αυτή η παρατήρηση επιτρέπει την αξιοποίηση των τεχνικών αραιής αναπαράστασης σήματος (compressive sensing), π.χ. [17]-[20], οι οποίες αναμένεται να οδηγήσουν σε αραιές εκτιμήσεις του διανύσματος των ποσοστών. Οι τεχνικές αυτές, προϋποθέτουν την ύπαρξη μιας φασματικής βιβλιοθήκης, η οποία συγκεντρώνει τις διαθέσιμες φασματικές υπογραφές πολλών διαφορετικών υλικών.

#### 1.2.4 Προτεινόμενη προσέγγιση κατά Bayes

Στην εργασία αυτή, παρουσιάζεται μία ιεραρχική προσέγγιση κατά Bayes για τον ημί-επιβλεπόμενο υπέρ-φασματικό διαχωρισμό, που αποτελεί ελαφρά παραλλαγή αυτής που αναπτύχθηκε αρχικά στο [1], και μοντελοποιεί τις ιδιότητες της αραιότητας και της μη αρνητικότητας του διανύσματος των ποσοστών, μέσω κατάλληλα ορισμένων παραμέτρων (τυχαίων μεταβλητών), στις οποίες ανατίθενται κατάλληλες εκ των προτέρων πυκνότητες πιθανότητας.

---

<sup>1</sup> Για συντομία, στη συνέχεια, θα αναφέρουμε απλώς «διάνυσμα ποσοστών».

Το μοντέλο μίξης, που υιοθετείται στην παρούσα εργασία (αλλά και σε πολλές άλλες περιπτώσεις υπερφασματικού διαχωρισμού), είναι το γραμμικό μοντέλο μίξης. Το μοντέλο αυτό, υποθέτει ότι το μετρούμενο φάσμα ενός μεικτού εικονοστοιχείου προκύπτει ως γραμμικός συνδυασμός των φασμάτων των υλικών που το συνθέτουν. Η μαθηματική έκφραση του LMM, που χρησιμοποιείται σε αυτήν την εργασία, καθώς και οι περιορισμοί, στους οποίους υπόκειται το εμπλεκόμενο στο μοντέλο διάνυσμα ποσοστών, διατυπώνονται πιο αυστηρά στο επόμενο κεφάλαιο.

Στο προτεινόμενο ιεραρχικό μοντέλο δύο επιπέδων, ανατίθενται κατάλληλες εκ των προτέρων κατανομές πιθανότητας (prior probability distribution ή prior) στις άγνωστες παραμέτρους, οι οποίες μοντελοποιούν την πρότερη γνώση σχετικά με τα φυσικά χαρακτηριστικά τους. Πιο συγκεκριμένα, για την μοντελοποίηση της μη αρνητικότητας των ποσοστών των συστατικών, υιοθετείται μία κολοβωμένη μη αρνητική κανονική κατανομή (truncated non-negative Gaussian distribution), ως πρώτο επίπεδο πρότερης γνώσης. Οι παράμετροι της διακύμανσης αυτής της κατανομής είναι, με τη σειρά τους, τυχαίες μεταβλητές, οι οποίες ακολουθούν εκθετική κατανομή. Αυτό, το δύο επιπέδων πιθανοτικό ιεραρχικό μοντέλο, ισοδυναμεί με μία Λαπλασιανή (Laplace) κατανομή για τα ποσοστά των υλικών, η οποία μοντελοποιεί την ιδιότητα της αραιότητας του διανύσματος των ποσοστών των υλικών, [21], [22]. Το προτεινόμενο αυτό ιεραρχικό μοντέλο, διατηρεί τη συζυγία (conjugacy) των κατανομών των παραμέτρων, μία ιδιότητα η οποία, στη συνέχεια, αξιοποιείται ώστε να προκύψουν εκφράσεις κλειστής μορφής για τις εκ των υστέρων κατανομές των εμπλεκόμενων στο μοντέλο παραμέτρων.

Όπως συμβαίνει συνήθως στη ανάλυση κατά Bayes, η προκύπτουσα από κοινού εκ των υστέρων κατανομή του προτεινόμενου ιεραρχικού μοντέλου, δεν έχει μία εύκολα διαχειρίσιμη αναλυτική μορφή. Για να ξεπεραστεί αυτό το εμπόδιο, αναπτύχθηκε ένας νέος επαναληπτικός αλγόριθμος, ο οποίος, για λόγους που θα γίνουν κατανοητοί στην συνέχεια, ονομάζεται BI-ICE-single. Ο αλγόριθμος αυτός, αποτελεί μία τροποποίηση του αλγορίθμου BI-ICE (Bayesian inference iterative conditional expectations), που παρουσιάστηκε στο [1].<sup>2</sup> Πιο συγκεκριμένα, όσον αφορά το διάνυσμα ποσοστών, τόσο ο BI-ICE-single, όσο και ο BI-ICE, υιοθετούν ένα αποτελεσματικό σχήμα, που ενημερώνει την εκ των υστέρων υπό συνθήκη μέση τιμή του διανύσματος των ποσοστών, ενώ όσον αφορά τις υπόλοιπες παραμέτρους που εμπλέκονται στο μοντέλο, οι εκ των υστέρων υπό συνθήκη μέσες τιμές αυτών ενημερώνονται, μέσω

---

<sup>2</sup> Η διαφορά των δύο αλγορίθμων θα γίνει ξεκάθαρη στα κεφάλαια που ακολουθούν.

απλών εκφράσεων κλειστής μορφής. Οι δύο αλγόριθμοι, σε κάθε επανάληψη, ενημερώνουν την εκ των υστέρων υπό συνθήκη μέση τιμή μιας παραμέτρου, με βάση τις τρέχουσες εκτιμήσεις των μέσων τιμών των υπολοίπων παραμέτρων.

Οι δύο προηγούμενοι αλγόριθμοι παράγουν μία αραιή εκτίμηση του διανύσματος ποσοστών, και ικανοποιούν τον περιορισμό της μη αρνητικότητας. Επίσης, παρέχουν και εκτιμήσεις όλων των άλλων παραμέτρων, που υπεισέρχονται στο πρόβλημα, μεταξύ των οποίων είναι η διακύμανση του προσθετικού θορύβου, που αλλοιώνει την υπερφασματική εικόνα. Ο BI-ICE-single είναι υπολογιστικά αποδοτικός, αφού εκτελώντας ελάχιστες επαναλήψεις καταλήγει σε εκτίμηση για τα ποσοστά των υλικών που συνθέτουν το εκάστοτε εικονοστοιχείο, εντοπίζοντας ποιά από αυτά είναι παρόντα και ποιά όχι (ποσοστά κοντά στο μηδέν δείχνουν ότι τα αντίστοιχα υλικά δεν συμμετέχουν στη σύνθεση του υπό εξέταση εικονοστοιχείου). Θα πρέπει να σημειωθεί, όμως, ότι ο αλγόριθμος BI-ICE-single παρουσιάζει υποδεέστερη συμπεριφορά από τον BI-ICE όταν οι φασματικές υπογραφές των υλικών παρουσιάζουν μεγάλο βαθμό συσχέτισης. Αντίθετα, οι δύο αλγόριθμοι παρουσιάζουν παρόμοια συμπεριφορά όταν ο βαθμός συσχέτισης των φασματικών υπογραφών των υλικών είναι μικρός.

Από την παραπάνω περιγραφή γίνεται σαφές ότι οι δύο αλγόριθμοι εξετάζουν το κάθε εικονοστοιχείο χωριστά, χωρίς να λαμβάνουν υπ' όψιν τυχόν συσχετίσεις μεταξύ γειτονικών εικονοστοιχείων (οι οποίες, όμως, αναμένονται να είναι αξιοσημείωτες).

Προκειμένου να αξιοποιηθεί η επιπλέον αυτή πληροφορία, προτείνεται μια νέα απλή τεχνική, στα πλαίσια της οποίας μπορούν να χρησιμοποιηθούν τόσο ο BI-ICE, όσο και ο BI-ICE-single, η οποία λαμβάνει υπ' όψιν της την πιθανή χωρική συσχέτιση μεταξύ των γειτονικών εικονοστοιχείων της υπερφασματικής εικόνας. Πειραματικά αποτελέσματα δείχνουν ότι, η νέα αυτή μέθοδος επεξεργασίας των εικονοστοιχείων οδηγεί σε ταχύτερη σύγκλιση και, κατά συνέπεια, καθιστά πιο αποδοτική υπολογιστικά την διαδικασία διαχωρισμού.

Το υπόλοιπο της παρούσας εργασίας οργανώνεται ως εξής:

Στο Κεφάλαιο 2, διατυπώνεται το πρόβλημα του αραιού ημί-επιβλεπόμενου φασματικού διαχωρισμού, και στο Κεφάλαιο 3 περιγράφεται το ιεραρχικό μοντέλο κατά Bayes που προτείνεται σε αυτήν την εργασία. Στο Κεφάλαιο 4, παρουσιάζεται ο προτεινόμενος αλγόριθμος συμπερασμού κατά Bayes, ο BI-ICE-single, ενώ στο Κεφάλαιο 5 περιγράφεται η τεχνική που χρησιμοποιεί την πιθανή χωρική συσχέτιση μεταξύ γειτονικών εικονοστοιχείων της υπερφασματικής εικόνας. Τα πειραματικά

αποτελέσματα, που προκύπτουν από τους αλγορίθμους BI-ICE και BI-ICE-single, τόσο σε δεδομένα προσομοίωσης όσο και σε πραγματικά δεδομένα, αναφέρονται στο Κεφάλαιο 6. Στη συνέχεια, στο Κεφάλαιο 7 παρατίθενται τα συμπεράσματα της όλης εργασίας. Επιπλέον, παρουσιάζεται ένας πίνακας ορολογίας αγγλικών και ελληνικών όρων, καθώς και τα αρκτικόλεξα που έχουν χρησιμοποιηθεί, ενώ στο Παράρτημα παρέχεται ο κώδικας MatLab που υλοποιήθηκε στα πλαίσια της παρούσας εργασίας. Τέλος, αναφέρονται όλες οι βιβλιογραφικές πηγές που χρησιμοποιήθηκαν για την εκπόνηση αυτής της εργασίας.

### 1.3 Συμβολισμοί

Στην παρούσα εργασία, για τις μαθηματικές εκφράσεις, χρησιμοποιούνται οι παρακάτω συμβολισμοί:

$(\cdot)^T \rightarrow$  αναστροφή.

$\|\cdot\|_1 \rightarrow \ell_1$  νόρμα, όπου  $\|\vec{x}\|_1 = \sum_{i=1}^N |x_i|$ , με  $\vec{x} = [x_1, x_2, \dots, x_N]^T$ .

$\|\cdot\|_2 \rightarrow \ell_2$  νόρμα, όπου  $\|\vec{x}\|_2^2 = \vec{x}^T \vec{x}$ .

$|\cdot| \rightarrow$  ορίζουσα ενός πίνακα ή απόλυτη τιμή.

$diag(\vec{x}) \rightarrow$  διαγώνιος πίνακας, ο οποίος έχει στη διαγώνιό του τα στοιχεία του διανύσματος  $\vec{x}$ .

$R^N \rightarrow$  ο  $N$  - διάστατος Ευκλείδειος χώρος.

$\mathbf{0} \rightarrow$  μηδενικό διάνυσμα.

$\mathbf{1} \rightarrow$  διάνυσμα του οποίου όλες οι συνιστώσες ισούνται με τη μονάδα.

$I_K \rightarrow$  ο  $K \times K$  ταυτοτικός πίνακας.

**Σημείωση:** για διανύσματα χρησιμοποιούμε τον συμβολισμό  $\vec{x}$ , ενώ για πίνακες τον συμβολισμό  $\Phi$ , (κεφαλαία έντονα γράμματα).

## 2. ΤΟ ΠΡΟΒΛΗΜΑ ΤΟΥ ΑΡΑΙΟΥ ΗΜΙ-ΕΠΙΒΛΕΠΟΜΕΝΟΥ ΦΑΣΜΑΤΙΚΟΥ ΔΙΑΧΩΡΙΣΜΟΥ

Στο κεφάλαιο αυτό, θα διατυπωθεί το πρόβλημα του αραιού ημί-επιβλεπόμενου φασματικού διαχωρισμού. Πριν από αυτό, όμως, θα παρατεθούν κάποιοι ορισμοί, που είναι απαραίτητοι για την παρουσίαση του προβλήματος.

### 2.1 Ορισμοί

Στην παρούσα εργασία, με  $\vec{y}$ , συμβολίζεται το διάνυσμα  $M \times 1$  διαστάσεων, το οποίο αποτελεί τη φασματική υπογραφή ενός εικονοστοιχείου μίας υπερφασματικής εικόνας, όπου  $M$  ο αριθμός των φασματικών ζωνών (spectral bands). Επίσης,  $\Phi = [\vec{\varphi}_1, \vec{\varphi}_2, \dots, \vec{\varphi}_N]$  είναι ο  $M \times N$  πίνακας (φασματικών) υπογραφών (signature matrix) του προβλήματος,  $M > N$ , όπου το  $M$ -διάστατο διάνυσμα  $\vec{\varphi}_i$  αντιπροσωπεύει την φασματική υπογραφή του  $i$ -οστού υλικού και  $N$  είναι ο συνολικός αριθμός των διαφορετικών υλικών. Τέλος, με  $\vec{w} = [w_1, w_2, \dots, w_N]^T$  συμβολίζουμε το  $N$ -διάστατο διάνυσμα ποσοστών (abundance vector) που σχετίζεται με το  $\vec{y}$ , ενώ το  $w_i$  δηλώνει το ποσοστό συμμετοχής του  $\vec{\varphi}_i$  στον καθορισμό του  $\vec{y}$ .

Ο προσθετικός θόρυβος  $\vec{n}$  είναι ένα τυχαίο διάνυσμα (random vector), που θεωρούμε ότι ακολουθεί κανονική κατανομή με μηδενική μέση τιμή και με ανεξάρτητα και ομοιόμορφα κατανομημένα (independent identically distributed - i.i.d.) στοιχεία. Θα γράφουμε  $\vec{n} | \beta \sim N(\vec{n} | 0, \beta^{-1} \mathbf{I}_M)$ , όπου  $\beta$  είναι το αντίστροφο της διασποράς (ακρίβεια) του θορύβου.

### 2.2 Διατύπωση του προβλήματος

Στην εργασία αυτή, όπως και σε αυτή που περιγράφεται στο [1], έχει υιοθετηθεί το γραμμικό μοντέλο μίξης, στο οποίο οι ποσότητες που αναφέρθηκαν προηγουμένως, συσχετίζονται με την παρακάτω σχέση

$$\vec{y} = \Phi \vec{w} + \vec{n}. \quad (1)$$

Λόγω της φύσης του προβλήματος, το διάνυσμα ποσοστών θα πρέπει να ικανοποιεί τους δύο παρακάτω περιορισμούς

$$(\alpha) \ w_i \geq 0, i = 1, 2, \dots, N \quad \text{και} \quad (\beta) \ \sum_{i=1}^N w_i = 1. \quad (2)$$

Με βάση την παραπάνω υπόθεση, εισάγεται μία ημί-επιβλεπόμενη τεχνική υπερφασματικού διαχωρισμού, στην οποία ο πίνακας των συστατικών,  $\Phi$ , θεωρείται γνωστός εκ των προτέρων. Όπως αναφέρθηκε και προηγουμένως, κάθε στήλη του  $\Phi$  περιέχει την φασματική υπογραφή  $\vec{\phi}_i$  ενός υλικού, τα στοιχεία της οποίας είναι μη αρνητικά, αφού αντιπροσωπεύουν τις ανακλαστικές τιμές σε συγκεκριμένες φασματικές ζώνες. Ο πίνακας  $\Phi$  μπορεί, είτε να προκύψει από κάποια έτοιμη φασματική βιβλιοθήκη, είτε να καθοριστεί με τη χρήση κάποιας τεχνικής εξαγωγής (extraction technique) υλικών, π.χ. [10]. Ωστόσο, ο πραγματικός αριθμός των υλικών που συνθέτουν το φάσμα ενός εικονοστοιχείου, ο οποίος συμβολίζεται με  $\xi$ , είναι άγνωστος και μπορεί να διαφέρει από εικονοστοιχείο σε εικονοστοιχείο. Δεδομένου ότι ένα εικονοστοιχείο, λογικά, θα απεικονίζει λίγα μόνο διαφορετικά υλικά, από τα συνολικά διαθέσιμα που υπάρχουν σε μία βιβλιοθήκη, αναμένεται ότι θα είναι  $\xi \ll N$ . Συνεπώς, η υπόθεση της αραιότητας για τα διανύσματα ποσοστών των εικονοστοιχείων είναι (στο παρόν πλαίσιο) μια λογική υπόθεση, η οποία μπορεί να αξιοποιηθεί για την πιστότερη μοντελοποίηση του προβλήματος και αναμένεται να οδηγήσει στην επίτευξη καλύτερων ποιοτικά λύσεων.

Συνοψίζοντας, στην ημί-επιβλεπόμενη μέθοδο φασματικού διαχωρισμού, κύριο μέλημα αποτελεί η εκτίμηση του διανύσματος ποσοστών  $\vec{w}$  για κάθε εικονοστοιχείο της εικόνας, το οποίο θα είναι μη αρνητικό και αραιό, με τα  $\xi$  από τα  $N$  στοιχεία του μη μηδενικά.

Το πρόβλημα του αραιού ημί-επιβλεπόμενου φασματικού διαχωρισμού μπορεί να αντιμετωπιστεί χρησιμοποιώντας, είτε μία από τις πρόσφατες προτεινόμενες τεχνικές αραιής αναπαράστασης σημάτων (compressive sensing techniques), που επικεντρώνονται, όμως, μόνο στην ιδιότητα της αραιότητας π.χ. [17], [19], [20], [24], είτε μία από τις τεχνικές τετραγωνικού προγραμματισμού, π.χ. [16], που επιβάλλουν επιτυχώς τους περιορισμούς που δόθηκαν στην (2), χωρίς, όμως, να εκμεταλλεύονται την αραιότητα.

Στο επόμενο κεφάλαιο, θα παρουσιαστεί ένα ιεραρχικό μοντέλο κατά Bayes, παρόμοιο με αυτό που αναπτύχθηκε στο [1], το οποίο λαμβάνει υπόψη του και την αραιότητα και τον περιορισμό του προβλήματος για μη αρνητικά ποσοστά συμμετοχής των υλικών στη σύνθεση ενός εικονοστοιχείου.

Μετά την παρουσίαση του μοντέλου αυτού, θα περιγραφεί ένας νέος αλγόριθμος, που αποτελεί μία τροποποίηση του BI-ICE αλγορίθμου του [1], και είναι κατάλληλος για συμπερασμό κατά Bayes (Bayesian inference). Τέλος, με μία μικρή αλλαγή-προσθήκη



Ανάπτυξη και αξιολόγηση αλγορίθμων συμπερασμού κατά Bayes κατάλληλων για φασματικό διαχωρισμό

στο αρχικό πρόβλημα (1), θα ενσωματωθεί στο γραμμικό μοντέλο μίξης και ο προσθετικός περιορισμός που αφορά το διάνυσμα των ποσοστών.

### 3. ΙΕΡΑΡΧΙΚΟ ΜΟΝΤΕΛΟ ΚΑΤΑ BAYES

#### 3.1 Εισαγωγή

Στην ενότητα αυτή, θα γίνει αρχικά, μία ακροθιγής αναφορά στο ιεραρχικό μοντέλο κατά Bayes για ημί-επιβλεπόμενο υπερφασματικό διαχωρισμό που αναπτύχθηκε στο [1], για την εκτίμηση του διανύσματος ποσοστών  $\vec{w}$  της (1), λαμβάνοντας υπόψη τόσο την ιδιότητα της αραιότητας, όσο και τον περιορισμό της μη αρνητικότητας που δόθηκε στη (2α). Στο μοντέλο αυτό, βασίζεται και το ιεραρχικό μοντέλο που προτείνεται σε αυτήν την εργασία και το οποίο θα εκτεθεί στη συνέχεια.

Και στα δύο αυτά μοντέλα (προτεινόμενο και αυτό που αναπτύχθηκε στο [1]), όλες οι άγνωστες ποσότητες θεωρούνται ότι είναι τυχαίες μεταβλητές, κάθε μία από τις οποίες μοντελοποιείται από μία εκ των προτέρων κατανομή πιθανότητας. Πιο συγκεκριμένα, εξαιτίας της μη αρνητικότητας των στοιχείων του  $\vec{w}$ , χρησιμοποιείται μία κολοβωμένη μη αρνητική κανονική κατανομή για το  $\vec{w}$ . Παρακάτω, παρατίθεται ο ορισμός μιας κολοβωμένης συνάρτησης πυκνότητας πιθανότητας. Η έννοια αυτή θα φανεί χρήσιμη στη συνέχεια.

Ορισμός 1: Έστω  $\mathbf{R}^N$  ένα υποσύνολο του  $R^N$  ( $\mathbf{R}^N \subseteq R^N$ ) με θετικό μέτρο κατά Lebesgue.

Με  $p(\cdot|\vec{\zeta})$  συμβολίζεται μία  $N$  - διάστατη κατανομή, όπου  $\vec{\zeta}$  είναι το διάνυσμα παραμέτρων της, και με  $p_{R^N}(\cdot|\vec{\zeta})$  συμβολίζεται η κολοβωμένη συνάρτηση πυκνότητας πιθανότητας (truncated probability density function-pdf), που προκύπτει από την κολόβωση της  $p(\cdot|\vec{\zeta})$  στο  $\mathbf{R}^N$ . Επίσης, με  $\vec{x} \sim p_{R^N}(\vec{x}|\vec{\zeta})$  δηλώνεται μία τυχαία μεταβλητή, της οποίας η συνάρτηση πυκνότητας πιθανότητας είναι ανάλογη (proportional) του  $p(\vec{x}|\vec{\zeta})I_{R^N}(\vec{x})$ , όπου

$$I_{R^N}(\vec{x}) = \begin{cases} 1, & x \in R^N \\ 0, & x \notin R^N . \end{cases} \quad (3)$$

Το μοντέλο που προτείνεται στο [1], περιλαμβάνει πολλές διακριτές μεταξύ τους παραμέτρους, για να ελέγχουν το επίπεδο της αραιότητας του διανύσματος ποσοστών, σε αντίθεση με άλλα ιεραρχικά μοντέλα, [20], [24], [25], αλλά και του μοντέλου που θα περιγραφεί παρακάτω, τα οποία χρησιμοποιούν μία παράμετρο για τον έλεγχο του

επιπέδου της αραιότητας.<sup>3</sup> Έχει αποδειχθεί ([1]) ότι αυτή η επέκταση καθιστά το μοντέλο αυτό ισοδύναμο με μία εκδοχή του προσαρμοστικού least absolute shrinkage and selection operator (Lasso) κριτηρίου, που λαμβάνει υπ' όψιν του τον περιορισμό της μη αρνητικότητας ([26]), και του οποίου η λύση παρέχει μία συνεπή (consistent) εκτίμηση για το διάνυσμα των ποσοστών.

Επίσης, το ιεραρχικό μοντέλο του [1], όπως και το μοντέλο που θα περιγραφεί εδώ, διατηρούν τη συζυγία των κατανομών των παραμέτρων, μία ιδιότητα η οποία αξιοποιείται, προκειμένου να εκφραστούν σε κλειστή μορφή οι εκ των υστέρων κατανομές των παραμέτρων.

Στη συνέχεια, ακολουθεί η περιγραφή του ιεραρχικού μοντέλου που προτείνεται σε αυτήν την εργασία, το οποίο, όπως αναφέρθηκε και προηγουμένως, ακολουθεί (σε μεγάλο βαθμό) την προσέγγιση του ιεραρχικού Bayesian μοντέλου [1], με τη διαφορά ότι αυτό έχει μόνο μία υπέρ-παραμέτρο για τον έλεγχο του επιπέδου της αραιότητας. Η σύγκριση των δύο μοντέλων θα γίνει στη συνέχεια του παρόντος κεφαλαίου.

### 3.2 Ανάλυση του ιεραρχικού μοντέλου

Στις ενότητες που ακολουθούν, παρουσιάζεται το ιεραρχικό Bayesian μοντέλο, που προτείνεται στην παρούσα εργασία, για την εκτίμηση του διανύσματος ποσοστών.

#### 3.2.1 Συνάρτηση πιθανοφάνειας

Λαμβάνοντας υπόψη το μοντέλο που ορίστηκε στην (1) και την ιδιότητα του προσθετικού Gaussian θορύβου, η συνάρτηση πιθανοφάνειας (likelihood) του  $\vec{y}$  εκφράζεται ως ακολούθως

$$p(\vec{y} | \vec{w}, \beta) = N(\vec{y} | \Phi\vec{w}, \beta^{-1}\mathbf{I}_M) = (2\pi)^{-\frac{M}{2}} \beta^{\frac{M}{2}} \exp\left[-\frac{\beta}{2} \|\vec{y} - \Phi\vec{w}\|_2^2\right]. \quad (4)$$

#### 3.2.2 Εκ των προτέρων κατανομές των παραμέτρων

Οι ιδιότητες της αραιότητας και της μη αρνητικότητας του  $\vec{w}$  εισάγονται μέσω κατάλληλα επιλεγμένων εκ των προτέρων κατανομών πιθανότητας. Μία ευρέως χρησιμοποιούμενη εκ των προτέρων κατανομή πιθανότητας που ευνοεί την αραιότητα,

<sup>3</sup> Στη θεωρία κατά Bayes, οι παράμετροι των εκ των προτέρων κατανομών πιθανότητας, ονομάζονται υπέρ-παραμέτροι (hyperparameters).

[20], [21], [24], [25], [27], είναι η Λαπλασιανή συνάρτηση πυκνότητας πιθανότητας, μηδενικής μέσης τιμής, η οποία για ένα  $w_i$  ορίζεται ως εξής

$$L(w_i | \lambda) = \frac{\lambda}{2} \exp[-\lambda |w_i|], \quad (5)$$

όπου το  $\lambda \geq 0$  είναι το αντίστροφο της παραμέτρου που καθορίζει τη μορφή (shape parameter) της Λαπλασιανής κατανομής. Υποθέτοντας ότι οι παράμετροι  $w_i$  είναι ανεξάρτητες μεταξύ τους, η  $N$ -διάστατη εκ των προτέρων κατανομή πιθανότητας του διανύσματος  $\vec{w} = [w_1, w_2, \dots, w_N]^T$  γράφεται ως

$$L(\vec{w} | \lambda) = \prod_{i=1}^N L(w_i | \lambda) = \left(\frac{\lambda}{2}\right)^N \exp[-\lambda \|\vec{w}\|_1]. \quad (6)$$

Μπορεί να αποδειχθεί, [21], ότι με βάση την εκ των προτέρων Λαπλασιανή κατανομή, η μέγιστη εκ των υστέρων εκτίμηση (maximum a posteriori estimation-MAP) του  $\vec{w}$  δίνεται από την σχέση

$$\hat{w} = \arg \min_w \left\{ \frac{\beta}{2} \|\vec{y} - \Phi \vec{w}\|_2^2 + \lambda \|\vec{w}\|_1 \right\}, \quad (7)$$

η οποία είναι η λύση του Lasso κριτηρίου της [28]. Όμως, αν η εκ των προτέρων Λαπλασιανή κατανομή εφαρμοστεί άμεσα στο αραιό διάνυσμα  $\vec{w}$ , η ιδιότητα της συζυγίας δεν θα ισχύει όσον αφορά την Gaussian συνάρτηση πιθανοφάνειας, που δίνεται στην (4), και ως εκ τούτου, η εκ των υστέρων συνάρτηση πυκνότητας πιθανότητας του  $\vec{w}$  δεν θα μπορεί να εξαχθεί σε κλειστή μορφή.

Στο [29] επισημαίνεται ότι, μία βασική ιδιότητα της Λαπλασιανής κατανομής είναι πως μπορεί να εκφραστεί σαν μία σταθμισμένη μίξη κανονικών κατανομών, με τους παράγοντες στάθμισης να ακολουθούν εκθετική πυκνότητα πιθανότητας, μέσω της ακόλουθης σχέσης

$$\frac{\lambda}{2} \exp[-\lambda |w_i|] = \int_0^{+\infty} \frac{1}{\sqrt{2\pi s}} \exp\left[-\frac{w_i^2}{2s}\right] \frac{\lambda^2}{2} \exp\left[-\frac{\lambda^2 s}{2}\right] ds, \quad \lambda > 0. \quad (8)$$

Η (8) υποδεικνύει ότι η εκ των προτέρων Λαπλασιανή κατανομή ισοδυναμεί με ένα ιεραρχικό Bayesian μοντέλο δύο επιπέδων, στο οποίο το  $\vec{w}$  ακολουθεί κανονική κατανομή (1<sup>ο</sup> επίπεδο), η διακύμανση της οποίας ακολουθεί εκθετική κατανομή (2<sup>ο</sup> επίπεδο). Αυτό το ιεραρχικό μοντέλο, που είναι ένας τύπος σταθμισμένης μίξης

Gaussian κατανομών (Gaussian scale mixture) (GSM), [30], έχει χρησιμοποιηθεί στα [20], [21], [24], [25], [27], [31]. Το κύριο πλεονέκτημα της διατύπωσης αυτής είναι ότι διατηρεί τη συζυγία μεταξύ των εκ των προτέρων και εκ των υστέρων κατανομών των εμπλεκόμενων παραμέτρων.

Προκειμένου να εξασφαλιστεί η μη αρνητικότητα των συνιστωσών του διανύσματος  $\vec{w}$ , το μοντέλο αυτό χρησιμοποιεί μία κολοβωμένη κανονική κατανομή στο μη αρνητικό orthant<sup>4</sup> του  $R^N$ , σαν μία πρώτου επιπέδου εκ των προτέρων κατανομή για το  $\vec{w}$ . Υποθέτοντας ότι, όλα τα  $w_i$  είναι ανεξάρτητα και ομοιόμορφα κατανεμημένα και ότι τα  $\gamma_i$  (κανονικοποιημένα από το  $\beta$ ) είναι οι διακυμάνσεις των  $w_i$ , η εκ των προτέρων κατανομή του διανύσματος  $\vec{w}$  μπορεί αναλυτικά να εκφραστεί ως

$$\begin{aligned}
 p(\vec{w} | \vec{\gamma}, \beta) &= \prod_{i=1}^N \left[ N_{R_+^1}(w_i | 0, \frac{\gamma_i}{\beta}) \right] \\
 &= \prod_{i=1}^N \left[ 2(2\pi)^{-\frac{1}{2}} \beta^{\frac{1}{2}} \gamma_i^{-\frac{1}{2}} \exp\left[-\frac{\beta}{2} \frac{w_i^2}{\gamma_i}\right] I_{R_+^1}(w_i) \right] \\
 &= 2^N (2\pi)^{-\frac{N}{2}} \beta^{\frac{N}{2}} |\Lambda|^{\frac{1}{2}} \exp\left[-\frac{\beta}{2} \vec{w}^T \Lambda \vec{w}\right] I_{R_+^N}(\vec{w}) \\
 &= N_{R_+^N}(\vec{w} | 0, \beta^{-1} \Lambda^{-1}), \tag{9}
 \end{aligned}$$

όπου  $R_+^1$  είναι το σύνολο των μη αρνητικών πραγματικών αριθμών,  $R_+^N$  είναι το μη αρνητικό orthant του  $R^N$ ,  $N_{R_+^N}(\cdot)$  υποδηλώνει την  $N$ -διάστατη κολοβωμένη κανονική κατανομή στο  $R_+^N$ , σύμφωνα με τον Ορισμό 1,  $\vec{\gamma} = [\gamma_1, \gamma_2, \dots, \gamma_N]$  είναι το  $N \times 1$  διαστάσεων διάνυσμα που περιέχει τις υπέρ-παραμέτρους,  $\gamma_i \geq 0, i = 1, 2, \dots, N$  (διασπορές των αντίστοιχων  $w_i$ ) και  $\Lambda$  είναι ο  $N \times N$  διαγώνιος πίνακας, με  $\Lambda^{-1} = \text{diag}(\vec{\gamma})$ . Σημειώνεται ότι, η χρήση του  $\beta$ , ως παράμετρος κανονικοποίησης στην (9), εξασφαλίζει την ιδιότητα της μοναδικής κορυφής (unimodality) της εκ των υστέρων κατανομής του  $\vec{w}$ , [25], [31].

<sup>4</sup> Το μη αρνητικό orthant είναι η περιοχή του  $R^N$ , που περιέχει όλα τα σημεία, των οποίων όλες οι συνιστώσες είναι μη αρνητικές.

Για την παράμετρο  $\beta$ , που εμφανίζεται στη συνάρτηση πιθανοφάνειας (4), υιοθετείται μία Γάμμα (Gamma) εκ των προτέρων κατανομή, η οποία ορίζεται ως

$$p(\beta | \kappa, \theta) = \Gamma(\beta | \kappa, \theta) = \frac{\theta^\kappa}{\Gamma(\kappa)} \beta^{\kappa-1} \exp[-\theta\beta], \quad (10)$$

όπου  $\beta \geq 0$ ,  $\kappa$  είναι η παράμετρος που καθορίζει τη μορφή,  $\kappa \geq 0$ , και  $\theta$  είναι το αντίστροφο της παραμέτρου που καθορίζει την κλίμακα (scale parameter) της Γάμμα κατανομής,  $\theta \geq 0$ . Η μέση τιμή και η διακύμανση της κατανομής αυτής είναι

$$E[p(\beta | \kappa, \theta)] = \frac{\kappa}{\theta} \text{ και } \text{var}[p(\beta | \kappa, \theta)] = \frac{\kappa}{\theta^2}, \text{ αντίστοιχα.}$$

### 3.2.3 Εκ των προτέρων κατανομές των υπέρ-παραμέτρων

Έχοντας ορίσει την κολοβωμένη κανονική κατανομή για τα  $w_i$ , θα ορίσουμε τώρα την εκθετική κατανομή για τα  $\gamma_i$  στο πνεύμα της (8). Η εκθετική εκ των προτέρων κατανομή πιθανότητας για το  $\gamma_i$  ορίζεται ως

$$p(\gamma_i | \lambda) = \Gamma\left(\gamma_i | 1, \frac{\lambda}{2}\right) = \frac{\lambda}{2} \exp\left[-\frac{\lambda}{2} \gamma_i\right], \quad i = 1, 2, \dots, N, \quad (11)$$

όπου το  $\lambda \geq 0$  είναι μία υπέρ-παράμετρος, που ελέγχει το επίπεδο της αραιότητας. Αν αυτές οι εκ των προτέρων κατανομές για τα στοιχεία του  $\vec{\gamma}$  αξιοποιηθούν από κοινού με την (9), τότε, η εκ των προτέρων κατανομή του  $\vec{w}$  μπορεί να δοθεί από την παρακάτω σχέση

$$\begin{aligned} p(w | \lambda, \beta) &= \int p(w | \gamma, \beta) p(\gamma | \lambda) d\gamma \\ &= \prod_{i=1}^N \int_0^{+\infty} p(w_i | \gamma_i, \beta) p(\gamma_i | \lambda) d\gamma_i \\ &= (\beta\lambda)^{\frac{N}{2}} \exp\left[-\sqrt{\beta\lambda} \sum_{i=1}^N |w_i|\right] I_{R_+^N}(\vec{w}) \\ &= L(\vec{w} | \sqrt{\beta\lambda}) I_{R_+^N}(\vec{w}). \end{aligned} \quad (12)$$

Σε συμφωνία με τον Ορισμό 1, η παράσταση  $L(\vec{w} | \sqrt{\beta\lambda}) I_{R_+^N}(\vec{w})$  συμβολίζεται ως  $L_{R_+^N}(\vec{w} | \sqrt{\beta\lambda})$ , που αποτελεί μία κολοβωμένη Λαπλασιανή κατανομή στο  $R_+^N$ .

Όπως θα δείξουμε παρακάτω, ο MAP εκτιμητής του  $\vec{w}$ , που ακολουθεί την κολοβωμένη Laplace εκ των προτέρων πιθανότητα, συμπίπτει με την εκτίμηση του  $\vec{w}$  που προκύπτει από τη βελτιστοποίηση του κριτηρίου Lasso (όπου όλα τα  $a_i$  θα είναι ίσα με  $\sqrt{\beta\lambda}$ ).

Η διαφορά του παρόντος μοντέλου σε σχέση με αυτό που περιγράφεται στο [1] είναι ότι, ενώ στο δεύτερο υπάρχει ένα διακριτό  $\lambda_i$  που σχετίζεται με κάθε  $\gamma_i$ , στο παρόν μοντέλο υπάρχει μόνο ένα  $\lambda$  που σχετίζεται με όλα τα  $\gamma_i$ . Ο λόγος που στο μοντέλο [1] χρησιμοποιήθηκε ένα διάνυσμα υπέρ-παραμέτρων  $\vec{\lambda}$ , αντί μιας βαθμωτής παραμέτρου  $\lambda$  για όλα τα  $\gamma_i$ , ήταν για να σχηματιστεί ένα κατά Bayes ιεραρχικό μοντέλο, ανάλογο του προσαρμοστικού Lasso κριτηρίου, που προτείνεται στο [26].

Όπως αναφέρθηκε και πιο πάνω, στο δεύτερο επίπεδο του ιεραρχικού μοντέλου, τα στοιχεία του  $\vec{\gamma}$ , μοντελοποιούνται από  $N$  ανεξάρτητες εκ των προτέρων Γάμμα κατανομές, παραμετροποιημένες όλες με μία κοινή παράμετρο  $\lambda$ , ως ακολούθως<sup>5</sup>

$$p(\gamma_i | \lambda) = \Gamma\left(\gamma_i | 1, \frac{\lambda}{2}\right) = \frac{\lambda}{2} \exp\left[-\frac{\lambda}{2} \gamma_i\right], \quad i = 1, 2, \dots, N. \quad (13)$$

Με την παραδοχή ότι όλα τα  $\gamma_i$  είναι ανεξάρτητα μεταξύ τους, η από κοινού κατανομή του  $\vec{\gamma}$  μπορεί να γραφεί ως

$$\begin{aligned} p(\vec{\gamma} | \lambda) &= \prod_{i=1}^N \left[ \frac{\lambda}{2} \exp\left[-\frac{\lambda}{2} \gamma_i\right] \right] \\ &= \left(\frac{\lambda}{2}\right)^N \exp\left[-\frac{\lambda}{2} \sum_{i=1}^N \gamma_i\right]. \end{aligned} \quad (14)$$

Τα δύο πρώτα στάδια του Bayesian μοντέλου, που συνοψίζονται στις σχέσεις (9) και (14), δίνουν μία μη αρνητική κολοβωμένη εκ των προτέρων Λαπλασιανή κατανομή που προωθεί την έννοια της αραιότητας. Αυτή η εκ των προτέρων κατανομή μπορεί να

<sup>5</sup> Η παρακάτω εξίσωση είναι η εξ. (11), η οποία επαναλαμβάνεται εδώ για λόγους διευκόλυνσης της μελέτης.

εκφραστεί συναρτήσει του  $\lambda$  προβαίνοντας σε περιθωριοποίηση (marginalization) ως προς το διάνυσμα  $\vec{\gamma}$ .

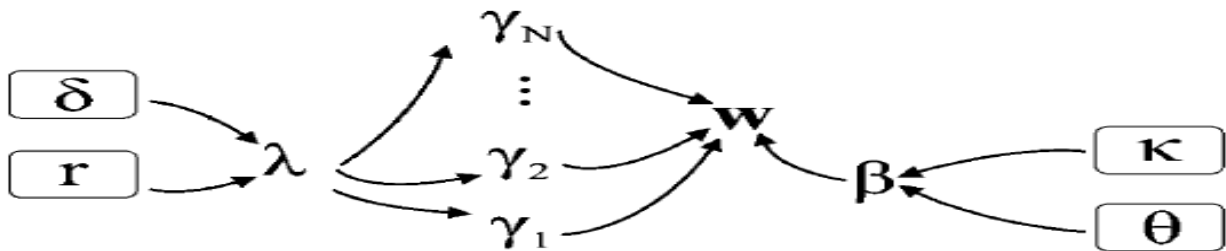
Συνήθως, η παράμετρος  $\lambda$  αντικατοπτρίζει μία πρότερη γνώση για το υπό εξέταση πρόβλημα και μπορεί, είτε, να οριστεί χειροκίνητα, είτε, να θεωρηθεί ως μία τυχαία μεταβλητή. Εδώ, επιλέγεται η δεύτερη εναλλακτική λύση, υποθέτοντας ότι η  $\lambda$  ακολουθεί μία εκ των προτέρων Γάμμα κατανομή, δηλαδή

$$p(\lambda | r, \delta) = \Gamma(\lambda | r, \delta) = \frac{\delta^r}{\Gamma(r)} \lambda^{r-1} \exp[-\delta\lambda], \quad (15)$$

όπου  $r$  και  $\delta$  είναι οι υπέρ-παράμετροι της κατανομής, με  $r, \delta \geq 0$ .

Και οι δύο εκ των προτέρων Γάμμα κατανομές που αναφέρθηκαν παραπάνω, δηλαδή του  $\beta$  στη (10) και του  $\lambda$  στη (15), είναι αρκετά ευέλικτες για να εκφράσουν τυχόν διαθέσιμη πρότερη πληροφορία, με κατάλληλη επιλογή των υπέρ - παραμέτρων τους.

Στο Σχήμα 1, παρουσιάζεται το προτεινόμενο ιεραρχικό Bayesian μοντέλο.



**Σχήμα 1:** Κατευθυνόμενο μη κυκλικό γράφημα του προτεινόμενου Bayesian μοντέλου. Οι παράμετροι που δεν είναι τυχαίες μεταβλητές εμφανίζονται σε ορθογώνια πλαίσια

Κλείνοντας το παρόν Κεφάλαιο, θα αποδείξουμε ότι το προτεινόμενο ιεραρχικό μοντέλο οδηγεί στην ίδια εκτίμηση για το  $\vec{w}$ , μ' αυτήν που προκύπτει από τη βελτιστοποίηση του κριτηρίου Lasso. Όπως φαίνεται στην εξ. (12), η κολοβωμένη εκ των προτέρων Λαπλασιανή κατανομή δίνεται από την σχέση

$$p(w | \lambda, \beta) = (\beta\lambda)^{\frac{N}{2}} \exp\left[-\sqrt{\beta\lambda} \sum_{i=1}^N |w_i|\right] I_{R_+^N}(\vec{w}). \quad (16)$$

Η μέγιστη εκ των υστέρων εκτίμηση του  $\vec{w}$  ορίζεται ως

$$\vec{w}_{MAP} = \arg \max_w p(\vec{w} | \vec{y}). \quad (17)$$



Από το θεώρημα του Bayes, η μέγιστη εκ των υστέρων εκτίμηση του  $\vec{w}$  μπορεί να εκφραστεί ως

$$\begin{aligned}\vec{w}_{MAP} &= \arg \max_w p(\vec{y} | \vec{w}, \beta) p(\vec{w} | \lambda, \beta) \\ &= \arg \min_w \{-\log [p(\vec{y} | \vec{w}, \beta) p(\vec{w} | \lambda, \beta)]\}.\end{aligned}\quad (18)$$

Αντικαθιστώντας στη (18) τη συνάρτηση πιθανοφάνειας από την (4) και την κολοβωμένη εκ των προτέρων Λαπλασιανή κατανομή από τη (16), η μέγιστη εκ των υστέρων εκτίμηση του  $\vec{w}$ , εκφράζεται ως

$$\begin{aligned}\vec{w}_{MAP} &= \arg \min_w \left\{ -\log \left[ (2\pi)^{-\frac{M}{2}} \beta^{\frac{M}{2}} \exp \left[ -\frac{\beta}{2} \|\vec{y} - \Phi \vec{w}\|_2^2 \right] (\beta\lambda)^{\frac{N}{2}} \exp \left[ -\sqrt{\beta\lambda} \sum_{i=1}^N |w_i| \right] I_{R_+^N}(\vec{w}) \right] \right\} \\ &= \arg \min_w \left[ \frac{\beta}{2} \|\vec{y} - \Phi \vec{w}\|_2^2 + \sqrt{\beta\lambda} \sum_{i=1}^N |w_i| - \log(I_{R_+^N}(\vec{w})) \right] \\ &= \arg \min_w \left[ \frac{\beta}{2} \|\vec{y} - \Phi \vec{w}\|_2^2 + \sqrt{\beta\lambda} \sum_{i=1}^N |w_i| \right],\end{aligned}\quad (19)$$

αφού  $-\log(I_{R_+^N}(\vec{w})) = \infty$  για  $\vec{w} \notin R_+^N$  και  $-\log(I_{R_+^N}(\vec{w})) = 0$  για  $\vec{w} \in R_+^N$ . Η παραπάνω ποσότητα εκφράζει το κριτήριο Lasso με συντελεστή ίσο με  $\sqrt{\beta\lambda}$ .

## 4. Ο ΠΡΟΤΕΙΝΟΜΕΝΟΣ ΑΛΓΟΡΙΘΜΟΣ ΣΥΜΠΕΡΑΣΜΟΥ ΚΑΤΑ BAYES

### 4.1 Εισαγωγή

Όπως συνηθίζεται στο συμπερασμό κατά Bayes (Bayesian inference), οι εκτιμήσεις των παραμέτρων (τυχαίων μεταβλητών), βασίζονται στις από κοινού εκ των υστέρων κατανομές τους.

Η από κοινού εκ των υστέρων κατανομή των παραμέτρων του μοντέλου που παρουσιάστηκε στο προηγούμενο κεφάλαιο, εκφράζεται ως ακολούθως

$$p(\vec{w}, \beta, \vec{\gamma}, \lambda | \vec{y}) = \frac{p(\vec{y} | \vec{w}, \beta) p(\vec{w} | \beta, \vec{\gamma}) p(\vec{\gamma} | \lambda) p(\lambda) p(\beta)}{p(y)}. \quad (20)$$

Όμως, η κατανομή αυτή δεν είναι εύκολα διαχειρίσιμη, αφού το ολοκλήρωμα

$$p(y) = \int \int \int \int p(\vec{y}, \vec{w}, \beta, \vec{\gamma}, \lambda) d\vec{w} d\vec{\gamma} d\lambda d\beta \quad (21)$$

δεν μπορεί να εκφραστεί σε κλειστή μορφή. Σε τέτοιες περιπτώσεις, ο δειγματολήπτης Gibbs [23] αποτελεί μία δημοφιλή εναλλακτική μέθοδο που ξεπερνάει το εμπόδιο αυτό. Συγκεκριμένα, ο δειγματολήπτης Gibbs παράγει τυχαία δείγματα, επαναληπτικά, από τις υπό συνθήκη εκ των υστέρων κατανομές των παραμέτρων του μοντέλου.

Στη συνέχεια του κεφαλαίου, θα εξαχθούν, οι υπό συνθήκη εκ των υστέρων κατανομές των παραμέτρων του μοντέλου και ακολούθως, οι αντίστοιχες εκ των υστέρων υπό συνθήκη μέσες τιμές (posterior conditional expectation). Με βάση αυτές τις μέσες τιμές (και σε αντίθεση με τα τυχαία δείγματα που χρησιμοποιούνται στο δειγματολήπτη Gibbs), θα αναπτυχθεί ο BI-ICE-single αλγόριθμος, ο οποίος οδηγεί σε αραιές εκτιμήσεις για το διάνυσμα ποσοστών, ικανοποιώντας τον περιορισμό της μη αρνητικότητας. Στο τέλος του κεφαλαίου, ενσωματώνεται και ο προσθετικός περιορισμός στο γραμμικό μοντέλο μίξης.

### 4.2 Υπό συνθήκη εκ των υστέρων κατανομές

Στο εδάφιο αυτό, θα οριστούν οι υπό συνθήκη εκ των υστέρων κατανομές των παραμέτρων του μοντέλου  $\vec{w}$ ,  $\vec{\gamma}$ ,  $\lambda$  και  $\beta$ .

Ξεκινώντας με το  $\vec{w}$ , μπορεί να αποδειχθεί, χρησιμοποιώντας τις σχέσεις (4) και (9), ότι η εκ των υστέρων υπό συνθήκη πυκνότητα πιθανότητάς του, μοντελοποιείται από μία κολοβωμένη πολυδιάστατη Gaussian κατανομή στο  $R_+^N$ , δηλαδή

$$\begin{aligned}
 p(\bar{w} | \bar{y}, \bar{\gamma}, \lambda, \beta) &= \frac{p(\bar{y} | \bar{w}, \beta) p(\bar{w} | \beta, \bar{\gamma})}{\int p(\bar{y} | \bar{w}, \beta) p(\bar{w} | \beta, \bar{\gamma}) d\bar{w}} \\
 &= N_{R_+^N}(\bar{w} | \bar{\mu}, \Sigma), \tag{22}
 \end{aligned}$$

όπου τα  $\bar{\mu}$  και  $\Sigma$  εκφράζονται ως ακολούθως, [32, θεώρημα 10.3],

$$\bar{\mu} = \beta \Sigma \Phi^T \bar{y} \tag{23}$$

$$\Sigma = \beta^{-1} [\Phi^T \Phi + \Lambda]^{-1}. \tag{24}$$

Όσον αναφορά την εκ των υστέρων υπό συνθήκη κατανομή για την παράμετρο  $\beta$ , μετά την εξάλειψη των όρων που είναι ανεξάρτητων του  $\beta$ , αυτή γράφεται ως

$$p(\beta | \bar{y}, \bar{w}, \bar{\gamma}, \lambda) = \frac{p(\bar{y} | \bar{w}, \beta) p(\bar{w} | \beta, \bar{\gamma}) p(\beta)}{\int_0^\infty p(\bar{y} | \bar{w}, \beta) p(\bar{w} | \beta, \bar{\gamma}) p(\beta) d\beta}. \tag{25}$$

Χρησιμοποιώντας τις σχέσεις (4), (9) και (10), αποδεικνύεται ότι η παράμετρος  $\beta$  μοντελοποιείται από μία Γάμμα κατανομή, που ορίζεται ως

$$p(\beta | \bar{y}, \bar{w}, \bar{\gamma}, \lambda) = \Gamma(\beta \left| \frac{M+N}{2} + \kappa, \frac{1}{2} \|\bar{y} - \Phi \bar{w}\|_2^2 + \theta + \frac{1}{2} \bar{w}^T \Lambda \bar{w} \right.). \tag{26}$$

Η υπό συνθήκη συνάρτηση πυκνότητας πιθανότητας της παραμέτρου  $\gamma_i$  μοντελοποιείται από μία γενικευμένη αντίστροφη (generalized inverse) Gaussian κατανομή, η οποία, χρησιμοποιώντας τις σχέσεις (9) και (13), υπολογίζεται ως

$$\begin{aligned}
 p(\gamma_i | \bar{y}, w_i, \lambda, \beta) &= \frac{p(\bar{y} | w_i, \beta) p(w_i | \gamma_i, \beta) p(\gamma_i | \lambda) p(\lambda) p(\beta)}{\int p(\bar{y} | w_i, \beta) p(w_i | \gamma_i, \beta) p(\gamma_i | \lambda) p(\lambda) p(\beta) d\gamma_i} \\
 &= \frac{p(w_i | \gamma_i, \beta) p(\gamma_i | \lambda)}{\int p(w_i | \gamma_i, \beta) p(\gamma_i | \lambda) d\gamma_i} \\
 &= \frac{2(2\pi)^{-\frac{1}{2}} \beta^{\frac{1}{2}} \gamma_i^{-\frac{1}{2}} \exp\left[-\frac{\beta w_i^2}{2 \gamma_i}\right] I_{R_+^1}(w_i) \bullet \frac{\lambda}{2} \exp\left[-\frac{\lambda}{2} \gamma_i\right]}{\int_0^\infty 2(2\pi)^{-\frac{1}{2}} \beta^{\frac{1}{2}} \gamma_i^{-\frac{1}{2}} \exp\left[-\frac{\beta w_i^2}{2 \gamma_i}\right] I_{R_+^1}(w_i) \bullet \frac{\lambda}{2} \exp\left[-\frac{\lambda}{2} \gamma_i\right] d\gamma_i}
 \end{aligned}$$

$$\begin{aligned}
 &= \frac{\gamma_i^{-\frac{1}{2}} \exp\left[-\frac{\beta w_i^2}{2 \gamma_i} - \frac{\lambda}{2} \gamma_i\right]}{\int_0^{\infty} \gamma_i^{-\frac{1}{2}} \exp\left[-\frac{\beta w_i^2}{2 \gamma_i} - \frac{\lambda}{2} \gamma_i\right] d\gamma_i} \\
 &= \frac{\gamma_i^{-\frac{1}{2}} \exp\left[-\frac{\beta w_i^2}{2 \gamma_i} - \frac{\lambda}{2} \gamma_i\right]}{\sqrt{\frac{2\pi}{\lambda}} \exp\left[-\sqrt{\beta\lambda} w_i\right]} \\
 &= \left(\frac{\lambda}{2\pi}\right)^{\frac{1}{2}} \gamma_i^{-\frac{1}{2}} \exp\left[-\frac{\beta w_i^2}{2 \gamma_i} - \frac{\lambda}{2} \gamma_i + \sqrt{\beta\lambda} |w_i|\right], \quad i=1,2,\dots,N. \tag{27}
 \end{aligned}$$

Για τον υπολογισμό του ολοκληρώματος στον παρανομαστή χρησιμοποιήθηκε το [33, εξ. 3.471.15].

Τέλος, η εκ των υστέρων υπό συνθήκη κατανομή του  $\lambda$ , δεδομένων των  $\bar{y}$ ,  $\bar{w}$ ,  $\bar{\gamma}$  και  $\beta$ , εκφράζεται, με την βοήθεια των (14) και (15), ως ακολούθως

$$\begin{aligned}
 p(\lambda | \bar{y}, \bar{w}, \bar{\gamma}, \beta) &= \frac{p(\bar{\gamma} | \lambda) p(\lambda)}{\int_0^{\infty} p(\bar{\gamma} | \lambda) p(\lambda) d\lambda} \\
 &= \frac{\prod_{i=1}^N \left[ \frac{\lambda}{2} \exp\left[-\frac{\lambda}{2} \gamma_i\right] \right] \cdot \frac{\delta^r}{\Gamma(r)} \lambda^{r-1} \exp[-\delta\lambda]}{\int_0^{\infty} \prod_{i=1}^N \left[ \frac{\lambda}{2} \exp\left[-\frac{\lambda}{2} \gamma_i\right] \right] \cdot \frac{\delta^r}{\Gamma(r)} \lambda^{r-1} \exp[-\delta\lambda] d\lambda} \\
 &= \frac{\left(\frac{\lambda}{2}\right)^N \exp\left[-\frac{\lambda}{2} \sum_{i=1}^N \gamma_i\right] \cdot \lambda^{r-1} \exp[-\delta\lambda]}{\int_0^{\infty} \left(\frac{\lambda}{2}\right)^N \exp\left[-\frac{\lambda}{2} \sum_{i=1}^N \gamma_i\right] \cdot \lambda^{r-1} \exp[-\delta\lambda] d\lambda} \\
 &= \frac{\lambda^N \exp\left[-\lambda \sum_{i=1}^N \frac{\gamma_i}{2}\right] \cdot \lambda^{r-1} \exp[-\delta\lambda]}{\int_0^{\infty} \lambda^{N+r-1} \exp\left[-\lambda \left(\sum_{i=1}^N \frac{\gamma_i}{2} + \delta\right)\right] d\lambda} \\
 &= \frac{\lambda^{N+r-1} \exp\left[-\lambda \left(\sum_{i=1}^N \frac{\gamma_i}{2} + \delta\right)\right]}{\int_0^{\infty} \lambda^{N+r-1} \exp\left[-\lambda \left(\sum_{i=1}^N \frac{\gamma_i}{2} + \delta\right)\right] d\lambda}
 \end{aligned}$$

$$\begin{aligned}
 &= \frac{\left(\sum_{i=1}^N \frac{\gamma_i}{2} + \delta\right)^{N+r} \lambda^{N+r-1} \exp\left[-\lambda\left(\sum_{i=1}^N \frac{\gamma_i}{2} + \delta\right)\right]}{\left(\sum_{i=1}^N \frac{\gamma_i}{2} + \delta\right)^{N+r} \int_0^{\infty} \lambda^{N+r-1} \exp\left[-\lambda\left(\sum_{i=1}^N \frac{\gamma_i}{2} + \delta\right)\right] d\lambda} \\
 &= \frac{\left(\sum_{i=1}^N \frac{\gamma_i}{2} + \delta\right)^{N+r} \cdot \lambda^{N+r-1} \exp\left[-\lambda\left(\sum_{i=1}^N \frac{\gamma_i}{2} + \delta\right)\right]}{\Gamma(N+r)} \\
 &= \Gamma\left(\lambda \mid N+r, \sum_{i=1}^N \frac{\gamma_i}{2} + \delta\right), \quad i=1,2,\dots,N. \tag{28}
 \end{aligned}$$

Δηλαδή, η εκ των υστέρων υπό συνθήκη κατανομή του  $\lambda$ , μοντελοποιείται από μία Γάμμα κατανομή, όπου  $N+r$  είναι η παράμετρος που καθορίζει τη μορφή,  $N+r \geq 0$  και  $\sum_{i=1}^N \frac{\gamma_i}{2} + \delta$  είναι το αντίστροφο της παραμέτρου που καθορίζει την κλίμακα της Γάμμα κατανομής,  $\sum_{i=1}^N \frac{\gamma_i}{2} + \delta \geq 0$ .

### 4.3 Ο αλγόριθμος BI-ICE-single

Όπως αναφέρθηκε και προηγουμένως, ο BI-ICE-single αλγόριθμος, μπορεί να θεωρηθεί ως μία ντετερμινιστική προσέγγιση του δειγματολήπτη Gibbs, όπου αντί των τυχαίων δειγμάτων που προκύπτουν από τις υπό συνθήκη εκ των υστέρων κατανομές των παραμέτρων (βλ. εξ. (22), (26), (27), (28)), χρησιμοποιούνται οι αντίστοιχες εκ των υστέρων υπό συνθήκη μέσες τιμές των κατανομών των παραμέτρων. Ο BI-ICE-single παράγει μία αραιή εκτίμηση του διανύσματος ποσοστών, ικανοποιώντας τον περιορισμό της μη αρνητικότητας.

Ο BI-ICE-single αλγόριθμος, όπως και ο BI-ICE, ενημερώνει σε κάθε επανάληψη τις εκ των υστέρων υπό συνθήκη μέσες τιμές, με βάση τις τρέχουσες εκτιμήσεις των υπολοίπων. Έτσι, προκύπτει ένα επαναληπτικό σχήμα μεταξύ των εκ των υστέρων υπό συνθήκη μέσων τιμών των  $\vec{w}$ ,  $\beta$ ,  $\gamma_i$  και  $\lambda$ .

Ωστόσο, όπως αναφέρεται και στο προηγούμενο κεφάλαιο, ο BI-ICE-single διαφέρει από τον BI-ICE, καθώς στον δεύτερο υιοθετείται μία μεταβλητή  $\lambda_i$  για κάθε  $\gamma_i$  (δηλ. έχουμε  $N$  το πλήθος παραμέτρους  $\lambda_i$ ), ενώ στον BI-ICE-single υιοθετείται μία μεταβλητή  $\lambda$  που σχετίζεται με όλα τα  $\gamma_i$ .

Τόσο ο BI-ICE-single, όσο και ο BI-ICE, αναμένεται να συγκλίνουν πιο γρήγορα από τον αρχικό Gibbs δειγματολήπτη.

Στη συνέχεια, υπολογίζονται οι εκ των υστέρων υπό συνθήκη μέσες τιμές των παραμέτρων του μοντέλου.

#### 4.3.1 Υπό συνθήκη μέση τιμή της $p(\vec{w} | \vec{y}, \vec{\gamma}, \lambda, \beta)$

Όπως φαίνεται στην (22), η  $p(\vec{w} | \vec{y}, \vec{\gamma}, \lambda, \beta)$  είναι μία κολοβωμένη Gaussian κατανομή στο  $R_+^N$ . Από το [34] είναι γνωστό ότι, στη μονοδιάστατη περίπτωση η υπό συνθήκη μέση τιμή μιας τυχαίας μεταβλητής  $x$ , η οποία μοντελοποιείται από μία κολοβωμένη Gaussian κατανομή στο  $R_+^1$ , υπολογίζεται ως

$$x \sim N_{R_+^1}(x | \mu^*, \sigma^{*2}) \Rightarrow E[x] = \mu^* + \frac{\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{\mu^{*2}}{\sigma^{*2}}\right)}{1 - \frac{1}{2} \operatorname{erfc}\left(\frac{\mu^*}{\sqrt{2}\sigma^*}\right)} \sigma^*, \quad (29)$$

όπου  $\operatorname{erfc}(\cdot)$  είναι η συμπληρωματική συνάρτηση λάθους. Δυστυχώς, από όσο γνωρίζουμε, στην  $N$ -διάστατη περίπτωση, δεν υπάρχει ανάλογη έκφραση κλειστής μορφής. Όμως, όπως σημειώνεται στα [35] και [36], η κατανομή του  $i$ -οστού στοιχείου του  $\vec{w}$ , που εξαρτάται από τα υπόλοιπα στοιχεία  $\vec{w}_{-i} = [w_1, \dots, w_{i-1}, w_{i+1}, \dots, w_N]^T$ , μπορεί να εκφραστεί ως

$$w_i | \vec{w}_{-i} \sim N_{R_+^1}(w_i | \mu_i^*, \sigma_{ii}^*), \quad (30)$$

όπου

$$\mu_i^* = \mu_i + \vec{\sigma}_{-i}^T \mathbf{\Sigma}_{-i-i}^{-1} (\vec{w}_{-i} - \vec{\mu}_{-i}) \quad (31)$$

$$\sigma_{ii}^* = \sigma_{ii} - \vec{\sigma}_{-i}^T \mathbf{\Sigma}_{-i-i}^{-1} \vec{\sigma}_{-i}. \quad (32)$$

Ανακαλώντας το γεγονός ότι η εκ των υστέρων υπό συνθήκη πυκνότητα πιθανότητας του  $\vec{w}$  μοντελοποιείται από μία κολοβωμένη πολυδιάστατη Gaussian κατανομή στο  $R_+^N$ , δηλαδή  $\vec{w} \sim N_{R_+^N}(\vec{w} | \vec{\mu}, \mathbf{\Sigma})$ , τα  $\mu_i$ ,  $\sigma_{ii}$  αντιπροσωπεύουν το  $i$ -οστό και  $ii$ -οστό στοιχείο των  $\vec{\mu}$  και  $\mathbf{\Sigma}$ , αντίστοιχα, ο  $\mathbf{\Sigma}_{-i-i}$  είναι ένας  $(N-1) \times (N-1)$  πίνακας, που προκύπτει από την αφαίρεση της  $i$ -οστής γραμμής και  $i$ -οστής στήλης του  $\mathbf{\Sigma}$ , ενώ, τέλος,

το  $\vec{\sigma}_{-i}$  είναι το διάνυσμα  $(N-1) \times 1$  διαστάσεων της  $i$ -οστής στήλης του  $\Sigma$ , μετά την αφαίρεση του  $i$ -οστού στοιχείου. Με την αξιοποίηση της (29), και χρησιμοποιώντας τις (31) και (32), οι υπό συνθήκη μέσες τιμές όλων των τυχαίων μεταβλητών  $w_i | \vec{w}_{-i}$ ,  $i=1,2,\dots,N$  μπορούν να υπολογιστούν αναλυτικά.

Παρακάτω, παρατίθεται η επαναληπτική διαδικασία για τον υπολογισμό της μέσης τιμής της  $p(\vec{w} | \vec{y}, \vec{\gamma}, \lambda, \beta)$ , που προτάθηκε στο [1] και υιοθετείται και από τον BI-ICE-single. Συγκεκριμένα, αναφέρεται η  $j$ -οστή επανάληψη,  $j=1,2,\dots$ , αυτής της διαδικασίας.

$$\begin{aligned}
 1. \quad w_1^{(j)} &= E \left[ p \left( w_1 | w_2^{(j-1)}, w_3^{(j-1)}, \dots, w_N^{(j-1)} \right) \right] \\
 2. \quad w_2^{(j)} &= E \left[ p \left( w_2 | w_1^{(j)}, w_3^{(j-1)}, \dots, w_N^{(j-1)} \right) \right] \\
 &\quad \vdots \\
 N. \quad w_N^{(j)} &= E \left[ p \left( w_N | w_1^{(j)}, w_2^{(j)}, \dots, w_{N-1}^{(j)} \right) \right].
 \end{aligned} \tag{33}$$

Η διαδικασία αυτή επαναλαμβάνεται μέχρις ότου επιτευχθεί σύγκλιση. Τα πειραματικά αποτελέσματα δείχνουν ότι, το επαναληπτικό σχήμα της (33) συγκλίνει στη μέση τιμή της  $\vec{w} \sim N_{R^N}(\vec{w} | \vec{\mu}, \Sigma)$  μετά από πολύ λίγες επαναλήψεις.

#### 4.3.2 Υπό συνθήκη μέση τιμή της $p(\beta | \vec{y}, \vec{w}, \vec{\gamma}, \lambda)$

Η μέση τιμή της Γάμμα κατανομής που δίνεται στην (26) είναι

$$E[p(\beta | \vec{y}, \vec{w}, \vec{\gamma}, \lambda)] = \left( \frac{M+N}{2} + \kappa \right) \cdot \left( \frac{1}{2} \|\vec{y} - \Phi \vec{w}\|_2^2 + \theta + \frac{1}{2} \vec{w}^T \mathbf{\Lambda} \vec{w} \right)^{-1}. \tag{34}$$

#### 4.3.3 Υπό συνθήκη μέση τιμή της $p(\gamma_i | \vec{y}, w_i, \lambda, \beta)$

Η μέση τιμή της  $p(\gamma_i | \vec{y}, w_i, \lambda, \beta)$  είναι

$$\begin{aligned}
 E[p(\gamma_i | \vec{y}, w_i, \lambda, \beta)] &= \int_0^\infty \gamma_i p(\gamma_i | \vec{y}, w_i, \lambda, \beta) d\gamma_i \\
 &= \int_0^\infty \left( \frac{\lambda}{2\pi} \right)^{\frac{1}{2}} \gamma_i^{\frac{1}{2}} \exp \left[ -\frac{\beta w_i^2}{2 \gamma_i} - \frac{\lambda}{2} \gamma_i + \sqrt{\beta \lambda} |w_i| \right] d\gamma_i \\
 &= \left( \frac{\lambda}{2\pi} \right)^{\frac{1}{2}} \exp \left[ \sqrt{\beta \lambda} |w_i| \right] \int_0^\infty \gamma_i^{\frac{1}{2}} \exp \left[ -\frac{\beta w_i^2}{2 \gamma_i} - \frac{\lambda}{2} \gamma_i \right] d\gamma_i
 \end{aligned}$$

$$= \left( \frac{2\lambda}{\pi} \right)^{\frac{1}{2}} \left( \frac{\beta w_i^2}{\lambda_i} \right) \exp \left[ \sqrt{\beta\lambda} |w_i| \right] K_{\frac{3}{2}} \left( \sqrt{\beta\lambda} |w_i| \right), \quad (35)$$

όπου,  $K_\nu(\cdot)$  είναι η τροποποιημένη συνάρτηση Bessel δεύτερου είδους τάξεως  $\nu$ . Για τον υπολογισμό του παραπάνω ολοκληρώματος χρησιμοποιήθηκε το [33, εξ. 3.471.9]. Επίσης, θεωρήσαμε ότι  $p(\gamma_i | \vec{y}, w_i, \lambda, \beta) = 0$ , για  $w_i < 0$ , γεγονός που δεν επηρεάζει τον BI-ICE-single αλγόριθμο, δεδομένου ότι έχει επιβληθεί ο περιορισμός της μη αρνητικότητας για όλα τα στοιχεία του  $\vec{w}$ .

#### 4.3.4 Υπό συνθήκη μέση τιμή της $p(\lambda | \vec{y}, w, \gamma, \beta)$

Η μέση τιμή της Γάμμα κατανομής, που δίνεται στην (28), εκφράζεται ως ακολούθως

$$E[p(\lambda | \vec{y}, w, \gamma, \beta)] = \frac{N+r}{\sum_{i=1}^N \frac{\gamma_i}{2} + \delta}. \quad (36)$$

#### 4.4 Παρουσίαση του BI-ICE-single αλγορίθμου

Ο BI-ICE-single αλγόριθμος, με βάση τις προηγούμενες εκφράσεις, συνοψίζεται στον Πίνακα 1. Όπως φαίνεται εκεί, ο αλγόριθμος αρχικοποιείται με  $\vec{\gamma}^{(0)} = 1$ ,  $\lambda = 1$  και  $\beta^{(0)} = 0.01 \|\vec{y}\|_2$  (όπως προτείνεται στο [24]).

Για την ενημέρωση της παραμέτρου  $\vec{w}^{(t)}$ , χρησιμοποιείται ένα βοηθητικό διάνυσμα, το  $\vec{v}$ . Το διάνυσμα αυτό, αρχικοποιείται με  $\vec{\mu}^{(t)}$  (η τιμή του  $\vec{\mu}$  στην  $t$  επανάληψη) και ενημερώνεται με την εκτέλεση μίας επανάληψης του σχήματος που περιγράφεται στην (33). Η προκύπτουσα τιμή του  $\vec{v}$  εκχωρείται στο  $\vec{w}^{(t)}$ . Αυτό, γίνεται με το σκεπτικό ότι, για διαγώνιο πίνακα  $\Sigma$  (που προκύπτει όταν οι στήλες του  $\Phi$  είναι ορθογώνιες), προκύπτει εύκολα από τις σχέσεις (31), (32), ότι τα  $w_i$  στην (33) είναι ασυσχέτιστα. Έτσι, μία επανάληψη είναι επαρκής για να παρθεί η μέση τιμή της  $N_{R^N}(\vec{w} | \vec{\mu}, \Sigma)$ . Αν και, αυτό δεν ισχύει όταν ο  $\Sigma$  δεν είναι διαγώνιος, τα πειραματικά αποτελέσματα δείχνουν ότι η εκτίμηση της μέσης τιμής της  $N_{R^N}(\vec{w} | \vec{\mu}, \Sigma)$ , που προκύπτει μετά την εκτέλεση μίας επανάληψης του σχήματος (33), είναι αρκετή στο πλαίσιο του BI-ICE-single αλγορίθμου.



Πειραματικά αποτελέσματα δείχνουν ότι ο BI-ICE-single οδηγεί σε αραιές εκτιμήσεις για το  $\vec{w}$ , όπου οι θέσεις των μη μηδενικών στοιχείων του  $\vec{w}$  προσδιορίζουν τα υλικά που συνθέτουν το φάσμα του υπό εξέταση εικονοστοιχείου, ενώ οι τιμές τους προσδιορίζουν το ποσοστό συμμετοχής των φασμάτων των υλικών στη σύνθεση του φάσματος του εικονοστοιχείου. Επιπλέον, όλες οι παράμετροι του μοντέλου εκτιμώνται με φυσικό τρόπο από τα δεδομένα, ως συνέπεια της Bayesian Lasso προσέγγισης που ακολουθείται σε αυτήν την εργασία. Αυτό βρίσκεται σε αντίθεση με τους ντετερμινιστικούς αλγορίθμους που επιλύουν το Lasso κριτήριο, π.χ. [17], [26], ή τις προσαρμοστικές μεθόδους, [37], που έχουν να αντιμετωπίσουν το πρόβλημα της κατάλληλης επιλογής (fine-tuning) των παραμέτρων (αντίστοιχων του  $\lambda$  στην περίπτωση μας), που ελέγχουν την αραιότητα της λύσης.

**Πίνακας 1: Ο αλγόριθμος BI-ICE-single**

Είσοδος  $\Phi, \vec{y}, \kappa, \theta, r, \delta$

Αρχικοποίησε  $\vec{\gamma}^{(0)} = 1, \lambda = 1, \beta^{(0)} = 0.01 \|\vec{y}\|_2$

for  $t=1, 2, \dots$  κάνε

– Υπολόγισε το  $\vec{w}^{(t)}$  όπως παρακάτω

Υπολόγισε τα  $\vec{\mu}^{(t)}, \Sigma^{(t)}$  με χρήση των (23), (24)

Θέσε  $\vec{v}^{(t)} = \vec{\mu}^{(t)}$

Υπολόγισε το  $u_1^{(t)} = E[p(u_1 | u_2^{(0)}, \dots, u_N^{(0)})]$  με χρήση των (31), (32) και (29)

Υπολόγισε το  $u_2^{(t)} = E[p(u_2 | u_1^{(1)}, u_3^{(0)}, \dots, u_N^{(0)})]$  με χρήση των (31), (32) και (29)

⋮

Υπολόγισε το  $u_N^{(t)} = E[p(u_N | u_1^{(1)}, u_2^{(1)}, \dots, u_{N-1}^{(1)})]$  με χρήση των (31), (32) και (29)

Θέσε  $\vec{w}^{(t)} = \vec{v}^{(t)}$

– Υπολόγισε το  $\beta^{(t)} = E[p(\beta | \vec{y}, \vec{w}^{(t)}, \vec{\gamma}^{(t-1)}, \lambda^{(t-1)})]$  με χρήση της (34)

– Υπολόγισε τα  $\gamma_i^{(t)} = E[p(\gamma_i | \vec{y}, w_i^{(t)}, \vec{\gamma}^{(t-1)}, \lambda^{(t-1)}, \beta^{(t)})]$ ,  $i=1, 2, \dots, N$ , με χρήση της (35)

– Υπολόγισε το  $\lambda^{(t)} = E[p(\lambda | \vec{y}, \vec{w}^{(t)}, \vec{\gamma}^{(t)}, \beta^{(t)})]$  με χρήση της (36)

endfor

Επιπρόσθετα, ο BI-ICE-single αλγόριθμος, εκτιμά τη διακύμανση του προσθετικού θορύβου του γραμμικού μοντέλου, όπως στο [13], καθώς και τη διακύμανση του διανύσματος ποσοστών. Η τελευταία εκτίμηση, σε συνδυασμό με την εκτίμηση του  $\vec{\mu}$ , παρέχει την εκ των υστέρων κατανομή του διανύσματος ποσοστών, η οποία μπορεί να

χρησιμοποιηθεί για την παραγωγή διαστημάτων εμπιστοσύνης, προκειμένου να αξιολογηθεί η αξιοπιστία του προτεινόμενου εκτιμητή.

Όσον αναφορά την πολυπλοκότητα του αλγορίθμου, από τον Πίνακα 1, γίνεται σαφές ότι, ο αλγόριθμος υπολογίζει μόνο εκφράσεις κλειστής μορφής. Η κύρια υπολογιστική επιβάρυνση, οφείλεται στον υπολογισμό των  $N$  αντίστροφων  $\Sigma_{-i-i}$ ,  $i=1,2,\dots,N$  πινάκων, που εμφανίζονται στις σχέσεις (31) και (32). Ωστόσο, στο [1, Παράρτημα Δ] αποδεικνύεται ότι, όλοι αυτοί οι αντίστροφοι πίνακες μπορούν να προκύψουν αποτελεσματικά από το  $\Sigma^{-1}$ . Έτσι, απαιτείται μόνο μία αντιστροφή πίνακα ανά επανάληψη (η οποία σχετίζεται με τον υπολογισμό του  $\Sigma$  στην (24)). Αυτό έχει σαν αποτέλεσμα την μείωση της πολυπλοκότητας του BI-ICE-single κατά μία τάξη μεγέθους ανά επανάληψη.

#### 4.5 Ενσωμάτωση προσθετικού περιορισμού στο γραμμικό μοντέλο μίξης

Το ιεραρχικό Bayesian μοντέλο, που παρουσιάστηκε στο προηγούμενο κεφάλαιο, λαμβάνει υπόψη του τη μη αρνητικότητα των στοιχείων του διανύσματος ποσοστών  $\vec{w}$ . Όμως, ο προσθετικός περιορισμός, σύμφωνα με τον οποίο όλα τα ποσοστά  $w_i$  πρέπει να αθροίζουν στη μονάδα, όπως αναφέρεται στη (2), δεν έχει εξεταστεί ακόμα. Όπως επισημαίνεται στο [38], ο περιορισμός αυτός δέχεται αυστηρές κριτικές. Στις πραγματικές υπερφασματικές εικόνες, οι φασματικές υπογραφές  $\vec{\varphi}_i$  ορίζονται, συνήθως, μέχρι ένα συντελεστή κλίμακας (up to a scale factor), και έτσι, ο προσθετικός περιορισμός πρέπει να αντικατασταθεί από ένα γενικευμένο περιορισμό (generalized constraint) της μορφής  $\sum c_i w_i = 1$ , όπου τα βάρη  $c_i$  δηλώνουν τους συντελεστές κλίμακας, οι οποίοι εξαρτώνται από το εκάστοτε εικονοστοιχείο. Επιπλέον, είναι γνωστό ότι η αραιή λύση ενός γραμμικού συστήματος, του οποίου το μητρώο  $\Phi$  έχει μη αρνητικά στοιχεία, οδηγεί σε λύση που ικανοποιεί τον γενικευμένο προσθετικό περιορισμό, [39]. Έτσι, μπορεί με βεβαιότητα να υποτεθεί ότι, ο αντίκτυπος της μη επιβολής του προσθετικού περιορισμού στην εκτέλεση του αλγορίθμου δεν αναμένεται να είναι σοβαρός. Παρ' όλα αυτά, στη συνέχεια περιγράφεται ένας αποτελεσματικός τρόπος για την επιβολή αυτού του περιορισμού, χρησιμοποιώντας όμως μία παράμετρο κανονικοποίησης.

Πρέπει να σημειωθεί ότι, η άμεση ενσωμάτωση του περιορισμού στο προτεινόμενο Bayesian μοντέλο θα απαιτούσε την κολόβωση της εκ των προτέρων κανονικής

κατανομής του  $\vec{w}$  σε ένα  $N$  - διάστατο μονόπλοκο (simplex), καθιστώντας αδύνατη την παραγωγή κλειστών μορφών εκφράσεων για τις εκ των υστέρων υπό συνθήκη κατανομές. Ένας τρόπος για να αντιμετωπιστεί αυτό το πρόβλημα είναι να επιβάλουμε τον προσθετικό περιορισμό ντετερμινιστικά (όπως προτείνεται και στα [1], [7], [14], [40, σελ. 586]), με την προσθήκη μίας επιπλέον εξίσωσης στο αρχικό γραμμικό μοντέλο μίξης, όπως παρακάτω

$$\begin{bmatrix} y \\ a \end{bmatrix} = \begin{bmatrix} \Phi \\ \alpha \mathbf{1}^T \end{bmatrix} \vec{w} + \begin{bmatrix} \vec{n} \\ 0 \end{bmatrix}, \quad (37)$$

όπου  $a$  είναι μία βαθμωτή παράμετρος (scalar parameter), η οποία καθορίζει την επίδραση του προσθετικού περιορισμού στην εκτίμηση του  $\vec{w}$ . Συγκεκριμένα, όσο μεγαλύτερη είναι η τιμή του  $a$ , τόσο πιο κοντά στη μονάδα θα αθροίζουν τα  $w_i$ . Θα πρέπει να επισημανθεί, ότι η προσθήκη της επιπλέον εξίσωσης στο LMM, (37), δεν επηρεάζει το προτεινόμενο ιεραρχικό Bayesian μοντέλο.

## 5. ΦΑΣΜΑΤΙΚΟΣ ΔΙΑΧΩΡΙΣΜΟΣ ΜΕ ΑΞΙΟΠΟΙΗΣΗ ΤΗΣ ΧΩΡΙΚΗΣ ΣΥΣΧΕΤΙΣΗΣ ΜΕΤΑΞΥ ΓΕΙΤΟΝΙΚΩΝ ΕΙΚΟΝΟΣΤΟΙΧΕΙΩΝ

Στο κεφάλαιο αυτό, θα περιγραφεί μια νέα απλή τεχνική για το φασματικό διαχωρισμό μίας υπερφασματικής εικόνας. Η τεχνική αυτή λαμβάνει υπ' όψιν της, την πιθανή χωρική συσχέτιση μεταξύ των γειτονικών εικονοστοιχείων της εικόνας. Πιο συγκεκριμένα, τα εικονοστοιχεία εξετάζονται το ένα μετά το άλλο κατά γραμμές, και (για καθένα από αυτά) λαμβάνει χώρα φασματικός διαχωρισμός, χρησιμοποιώντας είτε τον αλγόριθμο BI-ICE, είτε τον BI-ICE-single<sup>6</sup>. Όσον αφορά το πρώτο εικονοστοιχείο, η αρχικοποίηση του διάνυσματος ποσοστών, καθώς και των υπολοίπων παραμέτρων, γίνεται όπως περιγράφεται στο [24] για τον BI-ICE. Έπειτα, για κάθε επόμενο εικονοστοιχείο το διάνυσμα ποσοστών και οι υπόλοιπες παράμετροι αρχικοποιούνται, χρησιμοποιώντας τις αντίστοιχες τελικές τιμές που προέκυψαν μετά την ολοκλήρωση της επεξεργασίας του προηγούμενου εικονοστοιχείου. Στην περίπτωση που το προς εξέταση εικονοστοιχείο είναι το πρώτο μιας νέας γραμμής, η αρχικοποίηση γίνεται σύμφωνα με τις τελικές τιμές που προέκυψαν για το εικονοστοιχείο της προηγούμενης γραμμής, το οποίο βρίσκεται ακριβώς πάνω από το υπό εξέταση εικονοστοιχείο.

Η αξιοποίηση της πιθανής χωρικής συσχέτισης μεταξύ διαδοχικών εικονοστοιχείων αναμένεται να επιφέρει σημαντική εξοικονόμηση υπολογιστικής ισχύος κατά την εφαρμογή του φασματικού διαχωρισμού.

Η προτεινόμενη τεχνική, με χρήση του BI-ICE-single αλγορίθμου, συνοψίζεται στον Πίνακα 2. Σε αυτόν, οι παράμετροι που έχουν κάτω δεξιά το δείκτη *pr\_line*, χρησιμοποιούνται για την αποθήκευση της πληροφορίας που αφορά το πρώτο εικονοστοιχείο της προηγούμενης γραμμής. Επιπλέον, οι δείκτες που βρίσκονται πάνω δεξιά αναφέρονται στα δεδομένα της εκάστοτε παραμέτρου για την αντίστοιχη επανάληψη του αλγορίθμου, με τον δείκτη *end* να δηλώνει την τιμή της παραμέτρου μετά τη σύγκλιση. Τέλος, με **W** συμβολίζουμε έναν πίνακα διαστάσεων  $m \times n$ , όπου κάθε στοιχείο του αντιστοιχεί στο διάνυσμα των ποσοστών του αντίστοιχου εικονοστοιχείου, όπως αυτό προκύπτει μετά το πέρας της επεξεργασίας.

Η όλη διαδικασία δομείται γύρω από τον BI-ICE-single και χρησιμοποιούνται (όπου είναι απαραίτητο) οι αντίστοιχες παραδοχές που περιγράφονται στο [1], με τη μόνη διαφορά ότι γίνεται αξιοποίηση της πληροφορίας που προκύπτει από την επεξεργασία του

<sup>6</sup> Σημειώνεται ότι, θα μπορούσε να χρησιμοποιηθεί οποιοσδήποτε άλλος αλγόριθμος φασματικού διαχωρισμού στη θέση των δύο παραπάνω αλγορίθμων.

προηγούμενου από το υπό εξέταση εικονοστοιχείου της εικόνας. Αρχικά, η τεχνική αυτή δέχεται ως είσοδο τις παραμέτρους  $\kappa$ ,  $\theta$ ,  $r$  και  $\delta$ , όπως ακριβώς γίνεται και στον αλγόριθμο BI-ICE, καθώς και την εικόνα  $I$  διαστάσεων  $m \times n$ . Ακολουθεί η αρχικοποίηση των παραμέτρων με  $\vec{w}^{(0)} = 1, \vec{\gamma}_{pr\_line} = 1, \lambda_{pr\_line} = 1, \beta_{pr\_line} = 0.01 \|I_{1,1}\|_2$ .

**Πίνακας 2: Ο αλγόριθμος BI-ICE-single με αξιοποίηση της πληροφορίας της πιθανής χωρικής συσχέτισης**

```

Είσοδος  $\Phi, I, m, n, \kappa, \theta, r, \delta$ 

Αρχικοποίησε  $\vec{w}^{(0)} = 1, \vec{\gamma}_{pr\_line} = 1, \lambda_{pr\_line} = 1, \beta_{pr\_line} = 0.01 \|I_{1,1}\|_2$ 
for  $i=1, 2, \dots, n$  κάνε
    if  $i \neq 1$  τότε
         $\vec{w}^{(0)} = W_{i-1,1}$ 
    endif
     $\vec{\gamma}^{(0)} = \vec{\gamma}_{pr\_line}, \lambda^{(0)} = \lambda_{pr\_line}, \beta^{(0)} = \beta_{pr\_line}$ 
    for  $j=1, 2, \dots, m$  κάνε
         $\vec{y} = I_{i,j}$ 
        for  $t=1, 2, \dots$  κάνε
            -Υπολόγισε το  $\vec{w}^{(t)}$  όπως παρακάτω
                Υπολόγισε τα  $\vec{\mu}^{(t)}, \Sigma^{(t)}$  με χρήση των (23), (24)
                Θέσε  $\vec{v}^{(0)} = \vec{\mu}^{(t)}$ 
                Υπολόγισε το  $u_1^{(1)} = E[p(u_1 | u_2^{(0)}, \dots, u_N^{(0)})]$  με χρήση των (31), (32) και (29)
                Υπολόγισε το  $u_2^{(1)} = E[p(u_2 | u_1^{(1)}, u_3^{(0)}, \dots, u_N^{(0)})]$  με χρήση των (31), (32) και (29)
                ⋮
                Υπολόγισε το  $u_N^{(1)} = E[p(u_N | u_1^{(1)}, u_2^{(1)}, \dots, u_{N-1}^{(1)})]$  με χρήση των (31), (32) και (29)
                Θέσε  $\vec{w}^{(t)} = \vec{v}^{(1)}$ 
            -Υπολόγισε το  $\beta^{(t)} = E[p(\beta | \vec{y}, \vec{w}^{(t)}, \vec{\gamma}^{(t-1)}, \lambda^{(t-1)})]$  με χρήση της (34)
            -Υπολόγισε τα  $\gamma_i^{(t)} = E[p(\gamma_i | \vec{y}, w_i^{(t)}, \vec{\gamma}^{(t-1)}, \lambda^{(t-1)}, \beta^{(t)})]$ ,  $i = 1, 2, \dots, N$ , με χρήση της (35)
            -Υπολόγισε το  $\lambda^{(t)} = E[p(\lambda | \vec{y}, \vec{w}^{(t)}, \vec{\gamma}^{(t)}, \beta^{(t)})]$  με χρήση της (36)
        endfor
         $W_{i,j} = w^{(end)}, w^{(0)} = w^{(end)}, \vec{\gamma}^{(0)} = \vec{\gamma}^{(end)}, \lambda^{(0)} = \lambda^{(end)}, \beta^{(0)} = \beta^{(end)}$ 
        if  $j = 1$  τότε
             $\vec{\gamma}_{pr\_line} = \vec{\gamma}^{(end)}, \lambda_{pr\_line} = \lambda^{(end)}, \beta_{pr\_line} = \beta^{(end)}$ 
        endif
    endfor
endfor
    
```

Έπειτα, σε κάθε βήμα εκτελείται ο αλγόριθμος BI-ICE-single, όπου η αρχικοποίηση των παραμέτρων γίνεται χρησιμοποιώντας τις τιμές που προέκυψαν από τα προηγούμενα βήματα.

Στο επόμενο κεφάλαιο, που παρουσιάζονται τα πειραματικά αποτελέσματα για την προτεινόμενη μέθοδο, θα γίνει ξεκάθαρο ότι, όπως αναμενόταν, η χρήση της νέας αυτής τεχνικής επεξεργασίας των εικονοστοιχείων συγκλίνει πιο γρήγορα και καθιστά πιο αποδοτική την διαδικασία διαχωρισμού.

## 6. ΑΠΟΤΕΛΕΣΜΑΤΑ ΠΕΙΡΑΜΑΤΙΚΗΣ ΜΕΛΕΤΗΣ

### 6.1 Αποτελέσματα προσομοίωσης σε συνθετικά δεδομένα

Αυτή η ενότητα έχει ως στόχο να παρουσιάσει την αποτελεσματικότητα του αλγορίθμου BI-ICE-single, που παρουσιάστηκε στην παρούσα εργασία. Αυτό θα γίνει μέσω μιας σειράς πειραμάτων που σχετίζονται με το φασματικό διαχωρισμό μιας συνθετικής υπερφασματικής εικόνας. Ακολουθώντας τις πειραματικές ρυθμίσεις του [38], όπου γίνεται μια εμπειριστατωμένη σύγκριση αρκετών εκ των αλγορίθμων ημί-επιβλεπόμενου διαχωρισμού για αραιά δεδομένα, θεωρούμε τρία σύνολα φασματικών δεδομένων (βιβλιοθήκες φασματικών υπογραφών) για το προσομοιωμένο υπερφασματικό σκηνικό:

(α) το σύνολο  $\Phi \in R^{453 \times 220}$ , το οποίο είναι μητρώο που περιέχει τις φασματικές υπογραφές από 220 συστατικά, που έχουν επιλεγθεί από την φασματική βιβλιοθήκη USGS, [41],

(β) το σύνολο  $\Phi_1 \in R^{200 \times 100}$ , που είναι ένα μητρώο που περιέχει στοιχεία με ψευδοτυχαίες τιμές, τα οποία ακολουθούν την κανονική κατανομή, και

(γ) το σύνολο  $\Phi_2 \in R^{453 \times 220}$ , το οποίο είναι μητρώο με στοιχεία ανεξάρτητα και ομοιόμορφα κατανομημένα στο διάστημα [0 1].

Όπως αναμενόταν, οι φασματικές υπογραφές των υλικών του  $\Phi$  παρουσιάζουν πολύ υψηλό βαθμό συσχέτισης. Ο αριθμός συνθήκης (condition number) και η αμοιβαία συνάφεια (mutual coherence) του  $\Phi$  είναι  $36,182 \times 10^6$  και  $0,999933$ , αντίστοιχα. Οι αντίστοιχες τιμές για το μητρώο  $\Phi_1$  είναι  $5,4202$  και  $0,2344$ , και για το μητρώο  $\Phi_2$   $85,5764$  και  $0,8$ .

Σε όλες τις προσομοιώσεις, οι παρατηρήσεις θεωρείται ότι έχουν αλλοιωθεί από προσθετικό λευκό Gaussian θόρυβο (additive white Gaussian noise - AWGN). Η διασπορά του θορύβου αυτού, που υπάρχει σαν μεταβλητή στα πειράματα, καθορίζεται από την τιμή του λόγου σήματος-προς-θόρυβο (signal-to-noise ratio - SNR ή S/N).

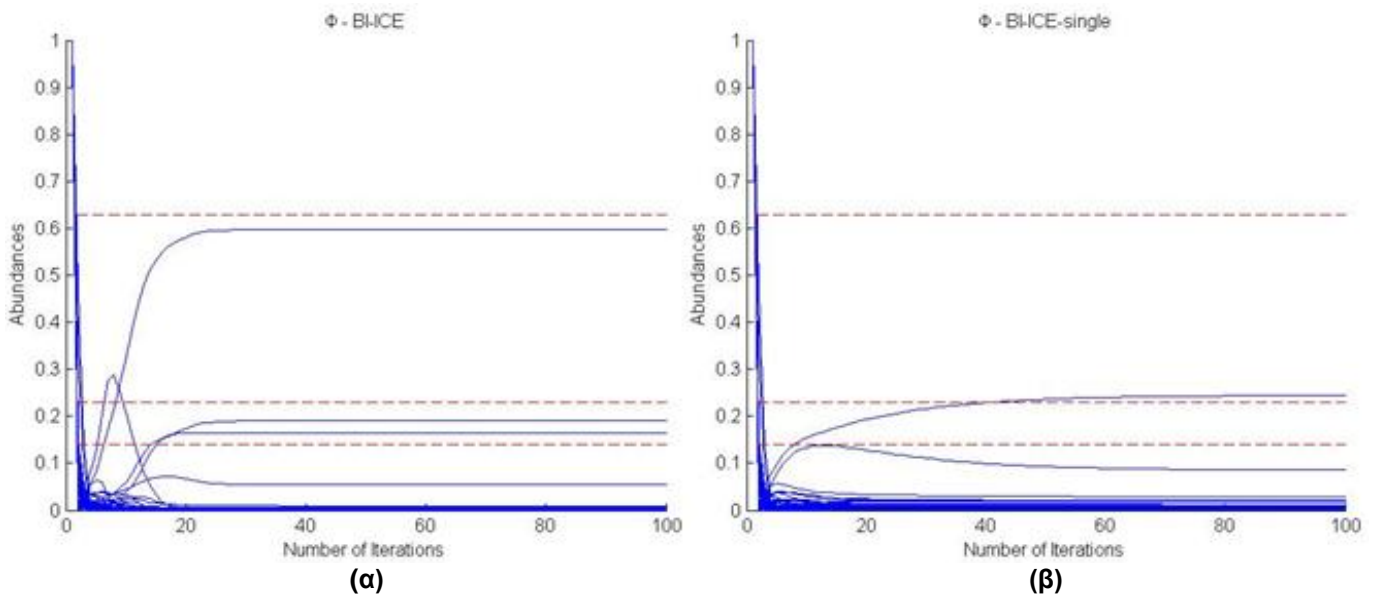
Ως μέτρο σύγκρισης για την αποτελεσματικότητα των αλγορίθμων χρησιμοποιείται το μέσο τετραγωνικό σφάλμα (Mean Squared Error - MSE), το οποίο δίνεται από τη σχέση

$$MSE = E \left[ \frac{\|\vec{w} - \vec{w}'\|_2^2}{\|\vec{w}\|_2^2} \right], \quad (38)$$

όπου  $\vec{w}$  και  $\vec{w}'$  είναι η πραγματική και η εκτιμώμενη τιμή του διανύσματος ποσοστών, αντίστοιχα.

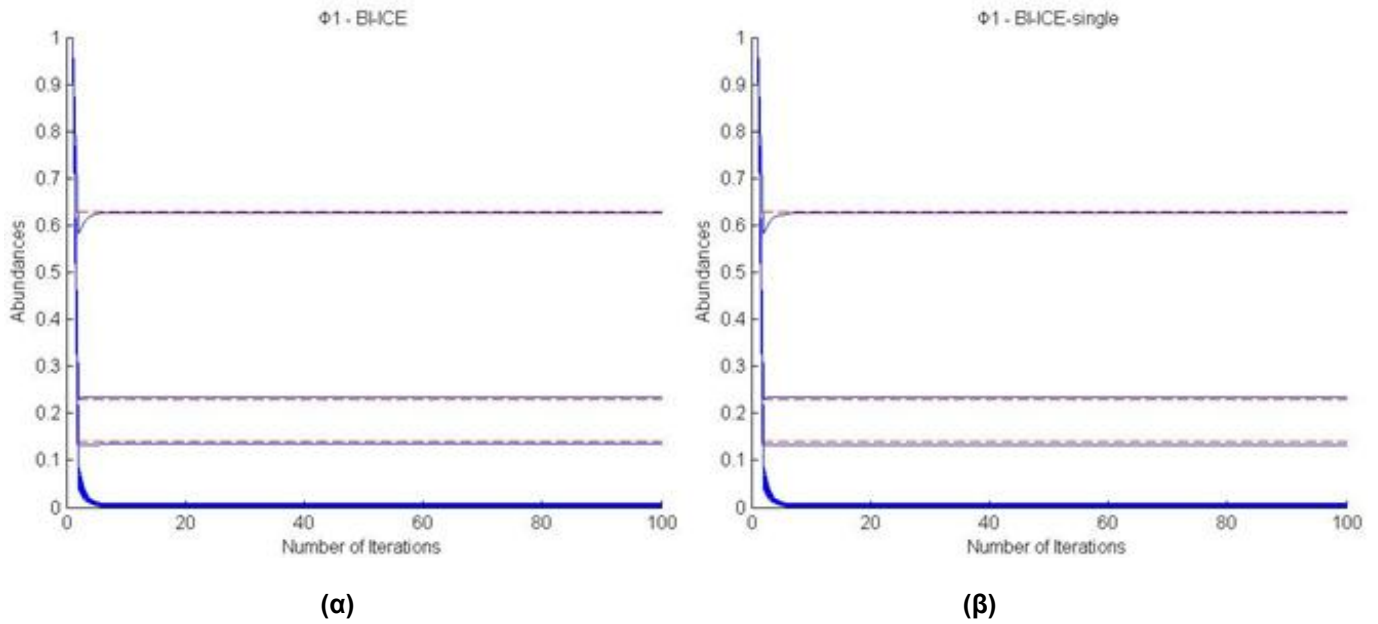
Τα πειράματα που λαμβάνουν χώρα στη συνέχεια αφορούν τον προτεινόμενο αλγόριθμο BI-ICE-single και τον αλγόριθμο BI-ICE ([1]), με σκοπό να εντοπιστούν οι πιθανές ομοιότητες, αλλά και οι διαφορές μεταξύ τους.

Αρχικά, στα Σχήματα 2, 3, 4, παρουσιάζεται τόσο η ταχεία σύγκλιση των αλγορίθμων όσο και οι αραιές εκτιμήσεις για το διάνυσμα ποσοστών, όταν (οι αλγόριθμοι) αυτοί εφαρμόζονται στα μητρώα  $\Phi$ ,  $\Phi_1$  και  $\Phi_2$ , αντίστοιχα. Στόχος αυτών των πειραμάτων είναι να γίνει σύγκριση των BI-ICE και BI-ICE-single ως προς την ταχύτητα σύγκλισης και την ακρίβεια εκτίμησης της τελικής τιμής του διανύσματος των ποσοστών. Στο πείραμα, έχει θεωρηθεί ένα εικονοστοιχείο του οποίου το φάσμα συντίθεται από τα φάσματα τριών υλικών, τα οποία συμμετέχουν σε ποσοστά 0.1397, 0.2305 και 0.6298, αντίστοιχα. Επιπλέον, υποτίθεται ότι το φάσμα του εικονοστοιχείου έχει αλλοιωθεί από λευκό θόρυβο, τέτοιον ώστε ο λόγος σήματος-προς-θόρυβο να είναι ίσος με 25 dB.

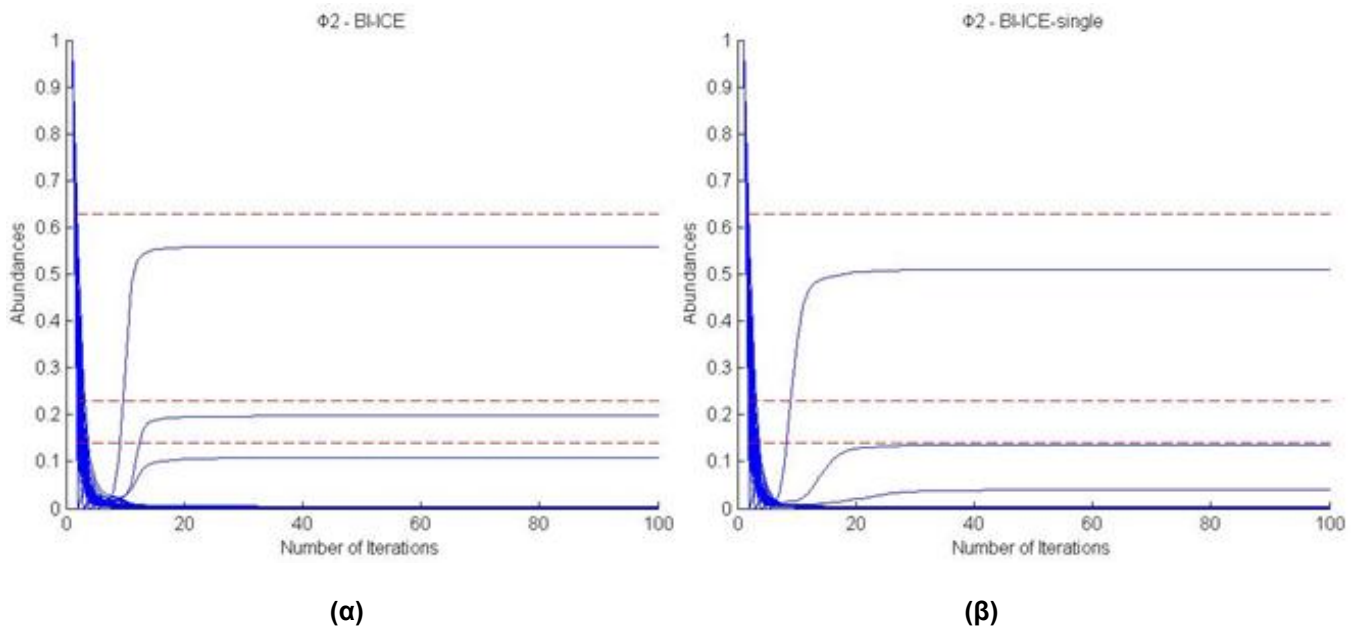


**Σχήμα 2:** Εκτίμηση των στοιχείων του διανύσματος ποσοστών  $\vec{w}$ , καθώς εξελίσσεται η εκτέλεση των αλγορίθμων (α) BI-ICE και (β) BI-ICE-single. Οι αλγόριθμοι εφαρμόζονται σε προσομοιωμένα δεδομένα, που έχουν δημιουργηθεί χρησιμοποιώντας ένα μητρώο φασματικών υπογραφών, οι οποίες παρουσιάζουν υψηλή συσχέτιση μεταξύ τους. Έχει προστεθεί λευκός θόρυβος (SNR = 25 dB). Διακεκομμένες γραμμές: Πραγματικές τιμές μη μηδενικών συνιστωσών. Συμπαγείς γραμμές: Εκτιμώμενες τιμές





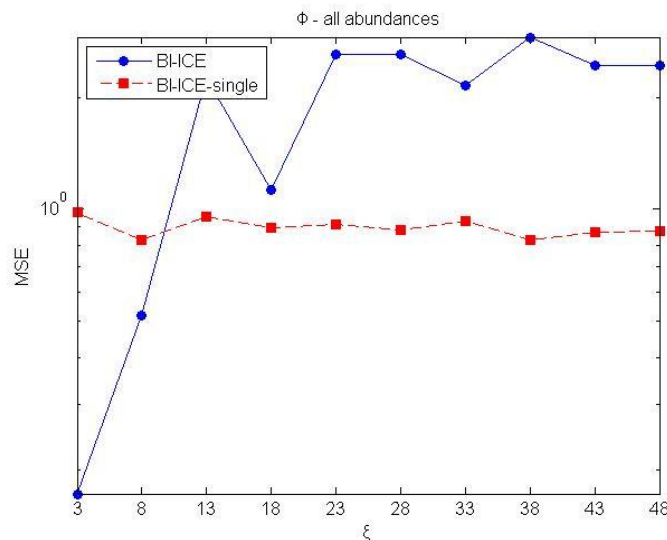
**Σχήμα 3:** Εκτίμηση των στοιχείων του διανύσματος ποσοστών  $\vec{w}$ , καθώς εξελίσσεται η εκτέλεση των αλγορίθμων (α) BI-ICE και (β) BI-ICE-single. Οι αλγόριθμοι εφαρμόζονται σε προσομοιωμένα δεδομένα, που έχουν δημιουργηθεί χρησιμοποιώντας ένα μητρώο φασματικών υπογραφών, οι οποίες ακολουθούν την κανονική κατανομή. Έχει προστεθεί λευκός θόρυβος (SNR = 25 dB). Διακεκομμένες γραμμές: Πραγματικές τιμές μη μηδενικών συνιστωσών. Συμπαγείς γραμμές: Εκτιμώμενες τιμές



**Σχήμα 4:** Εκτίμηση των στοιχείων του διανύσματος ποσοστών  $\vec{w}$ , καθώς εξελίσσεται η εκτέλεση των αλγορίθμων (α) BI-ICE και (β) BI-ICE-single. Οι αλγόριθμοι εφαρμόζονται σε προσομοιωμένα δεδομένα, που έχουν δημιουργηθεί χρησιμοποιώντας ένα μητρώο φασματικών υπογραφών, οι οποίες είναι ανεξάρτητα και ομοιόμορφα κατανομημένες. Έχει προστεθεί λευκός θόρυβος (SNR = 25 dB). Διακεκομμένες γραμμές: Πραγματικές τιμές μη μηδενικών συνιστωσών. Συμπαγείς γραμμές: Εκτιμώμενες τιμές

Παρατηρούμε ότι για το μητρώο  $\Phi$ , του οποίου οι στήλες παρουσιάζουν υψηλή συσχέτιση, ο αλγόριθμος BI-ICE συγκλίνει σε 15 περίπου επαναλήψεις, δίνοντας εκτιμήσεις των ποσοστών που είναι πολύ κοντά στις πραγματικές. Πιο συγκεκριμένα, φαίνεται να καθορίζει σωστά τα ποσοστά συμμετοχής των φασμάτων των υλικών που συνθέτουν την φασματική υπογραφή του εικονοστοιχείου, ενώ όλα τα υπόλοιπα ποσοστά συγκλίνουν στο μηδέν. Αντίθετα, ο BI-ICE-single χρειάζεται περισσότερες επαναλήψεις προκειμένου να συγκλίνει και, επιπλέον, οι μη μηδενικές τιμές ποσοστών στις οποίες συγκλίνει, δεν αποτελούν καλές εκτιμήσεις των αντίστοιχων πραγματικών. Παρόλα αυτά, ο αλγόριθμος φαίνεται ότι εντοπίζει σωστά τα μη μηδενικά στοιχεία.

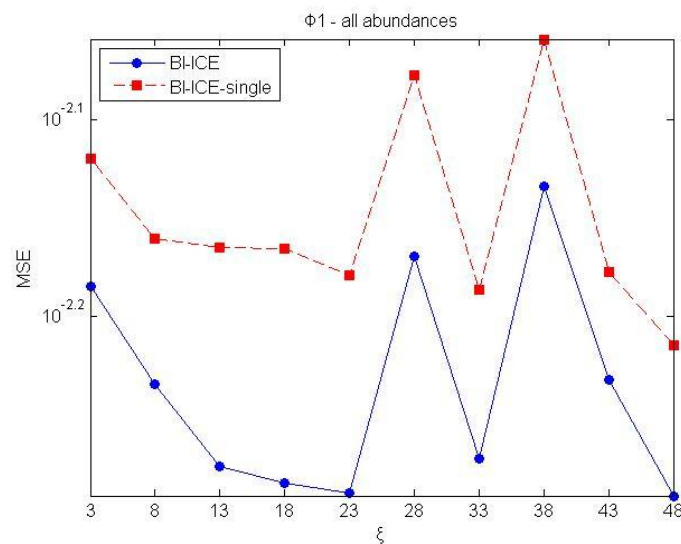
Σε αντίθεση με την παραπάνω περίπτωση, τα αποτελέσματα βελτιώνονται σημαντικά για μητρώα των οποίων οι στήλες παρουσιάζουν χαμηλή συσχέτιση, όπως είναι τα  $\Phi_1$  και  $\Phi_2$ . Πιο συγκεκριμένα, στα Σχήματα 3 και 4 φαίνεται ότι στις περιπτώσεις αυτές, η ταχύτητα σύγκλισης αυξάνεται σημαντικά. Επιπλέον, για το μητρώο  $\Phi_1$  (κανονική κατανομή), και οι δύο αλγόριθμοι δίνουν ικανοποιητικές εκτιμήσεις των μη μηδενικών ποσοστών. Ωστόσο, η απόδοση των αλγορίθμων είναι χειρότερη για το μητρώο  $\Phi_2$  (ομοιόμορφη κατανομή) σε σχέση με το  $\Phi_1$ . Τέλος, όσον αφορά το μητρώο  $\Phi_2$ , ο αλγόριθμος BI-ICE δίνει καλύτερες εκτιμήσεις των μη μηδενικών ποσοστών σε σχέση με τον BI-ICE-single.



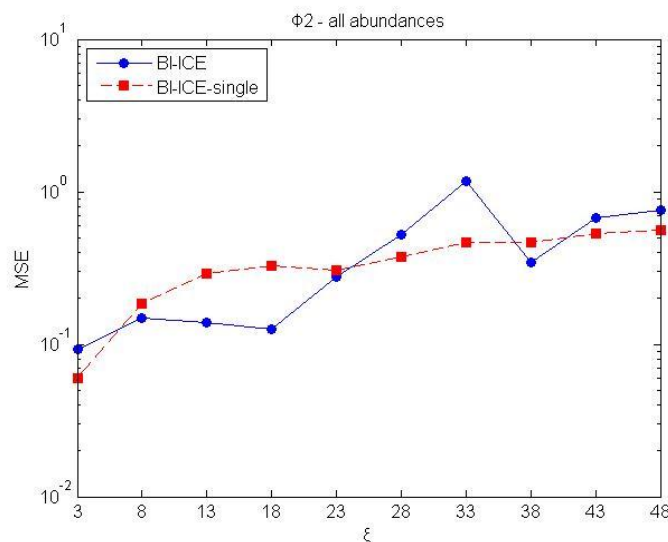
**Σχήμα 5:** Μέσο τετραγωνικό σφάλμα που παράγεται από τους BI-ICE και BI-ICE-single, για διαφόρους βαθμούς αραιότητας. Οι δύο αλγόριθμοι εφαρμόζονται σε προσομοιωμένα δεδομένα που προκύπτουν από το μητρώο  $\Phi$ , ενώ έχει προστεθεί λευκός θόρυβος (SNR = 20 dB)

Στα Σχήματα 5-7 παριστάνονται γραφικά οι αντίστοιχες καμπύλες του μέσου τετραγωνικού σφάλματος για τα τρία μητρώα και για διαφορετικούς βαθμούς αραιότητας που κυμαίνονται από 3 έως 48. Για κάθε βαθμό αραιότητας έχουν δημιουργηθεί δύο

πειραματικά εικονοστοιχεία, και οι τιμές του γραφήματος αποτελούν τις μέσες τιμές των σφαλμάτων που προέκυψαν μετά την επεξεργασία των εικονοστοιχείων αυτών. Το σφάλμα εκτιμάται σε σχέση με το πραγματικό διάνυσμα ποσοστών  $\vec{w}$ , λαμβάνοντας υπ' όψιν όλες τις συνιστώσες τόσο του  $\vec{w}$ , όσο και του εκτιμώμενου διανύσματος  $\vec{w}'$ . Σε όλες τις μετρήσεις έχει προστεθεί λευκός θόρυβος για τον οποίο ο λόγος σήματος-προς-θόρυβο είναι 20 dB.



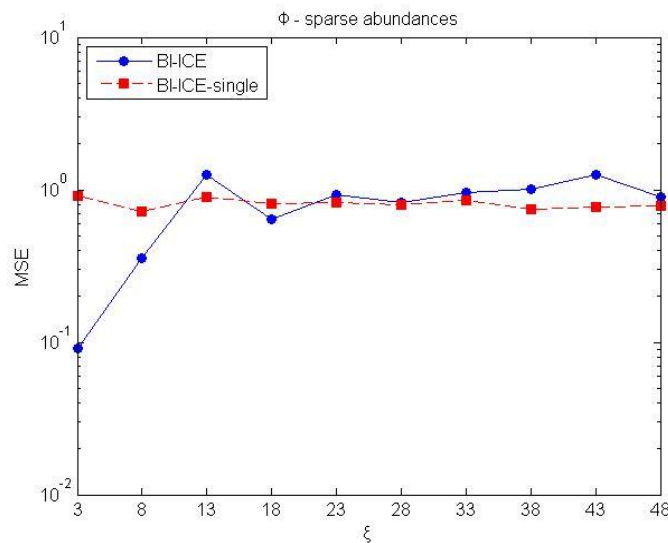
**Σχήμα 6:** Μέσο τετραγωνικό σφάλμα που παράγεται από τους BI-ICE και BI-ICE-single, για διαφόρους βαθμούς αραιότητας. Οι δύο αλγόριθμοι εφαρμόζονται σε προσομοιωμένα δεδομένα που προκύπτουν από το μητρώο  $\Phi_1$ , ενώ έχει προστεθεί λευκός θόρυβος (SNR = 20 dB)



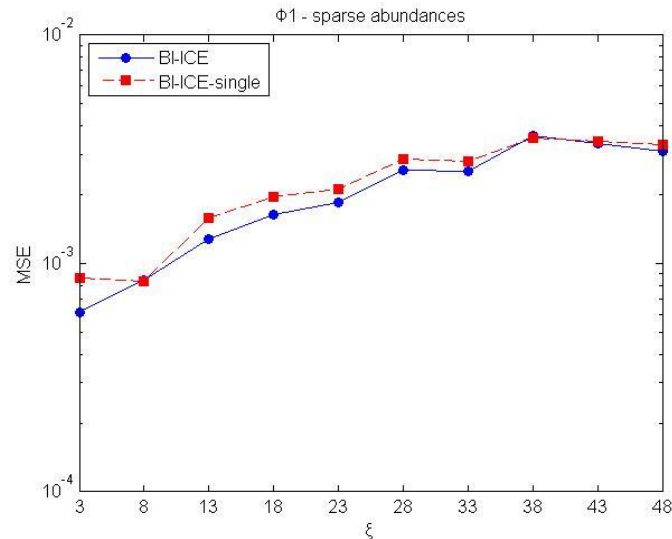
**Σχήμα 7:** Μέσο τετραγωνικό σφάλμα που παράγεται από τους BI-ICE και BI-ICE-single, για διαφόρους βαθμούς αραιότητας. Οι δύο αλγόριθμοι εφαρμόζονται σε προσομοιωμένα δεδομένα που προκύπτουν από το μητρώο  $\Phi_2$ , ενώ έχει προστεθεί λευκός θόρυβος (SNR = 20 dB)

Από τα σχήματα αυτά προκύπτει ότι οι δύο αλγόριθμοι έχουν παρόμοια συμπεριφορά. Ωστόσο, παρατηρούμε ότι καθώς ο βαθμός αραιότητας αυξάνεται, το MSE για τον BI-ICE-single δεν παρουσιάζει σημαντικές διακυμάνσεις για κανένα από τα μητρώα  $\Phi$ ,  $\Phi_1$ ,  $\Phi_2$ . Αντίθετα, ο BI-ICE παρουσιάζει σημαντικές διακυμάνσεις στην περίπτωση του μητρώου  $\Phi$ , καθώς αυξάνεται ο βαθμός αραιότητας. Αξίζει, ακόμα, να σημειωθεί ότι όσον αφορά τα μητρώα  $\Phi_1$  και  $\Phi_2$ , οι δύο αλγόριθμοι παρουσιάζουν παρόμοιες τιμές μέσου τετραγωνικού σφάλματος, για τους διάφορους βαθμούς αραιότητας.

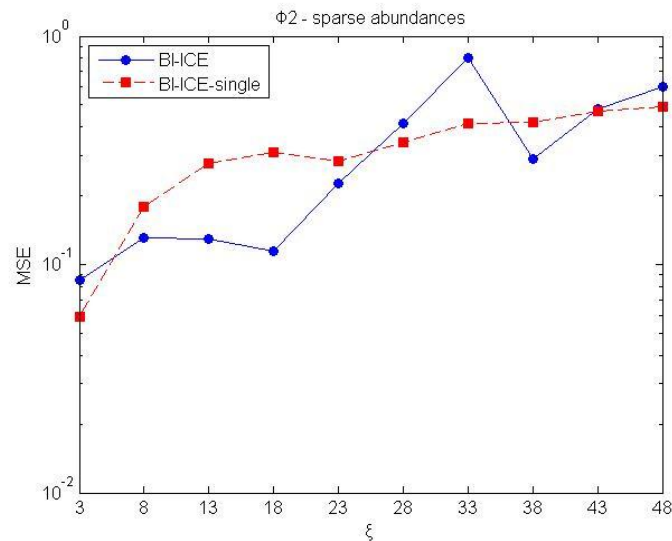
Θα ήταν όμως ενδιαφέρον, να εξετάσουμε τι συμβαίνει όταν, αντί να λαμβάνουμε υπ' όψιν όλες τις συνιστώσες του  $\vec{w}$ , λαμβάνουμε υπ' όψιν μόνο τις μη μηδενικές συνιστώσες του (οι οποίες προσδιορίζουν και τα υλικά που είναι παρόντα στο υπό εξέταση εικονοστοιχείο). Οι αντίστοιχες γραφικές παραστάσεις δίνονται στα Σχήματα 8-10, τα οποία είναι σε ένα προς ένα αντιστοιχία με τα Σχήματα 5-7, με τη μόνη διαφορά ότι το σφάλμα δεν αφορά όλο το διάνυσμα  $\vec{w}$ , αλλά υπολογίζεται λαμβάνοντας υπ' όψιν μόνο τις συνιστώσες που αντιστοιχούν στα υλικά που είναι παρόντα στο εικονοστοιχείο.



**Σχήμα 8:** Μέσο τετραγωνικό σφάλμα που προκύπτει λαμβάνοντας υπ' όψιν μόνο τα υλικά που συνθέτουν το εικονοστοιχείο, το οποίο παράγεται από τους BI-ICE και BI-ICE-single, για διάφορους βαθμούς αραιότητας. Οι δύο αλγόριθμοι εφαρμόζονται σε προσομοιωμένα δεδομένα που προκύπτουν από το μητρώο  $\Phi$ , ενώ έχει προστεθεί λευκός θόρυβος (SNR = 20 dB)



**Σχήμα 9:** Μέσο τετραγωνικό σφάλμα που προκύπτει λαμβάνοντας υπ' όψιν μόνο τα υλικά που συνθέτουν το εικονοστοιχείο, το οποίο παράγεται από τους BI-ICE και BI-ICE-single, για διαφόρους βαθμούς αραιότητας. Οι δύο αλγόριθμοι εφαρμόζονται σε προσομοιωμένα δεδομένα που προκύπτουν από το μητρώο  $\Phi_1$ , ενώ έχει προστεθεί λευκός θόρυβος (SNR = 20 dB)

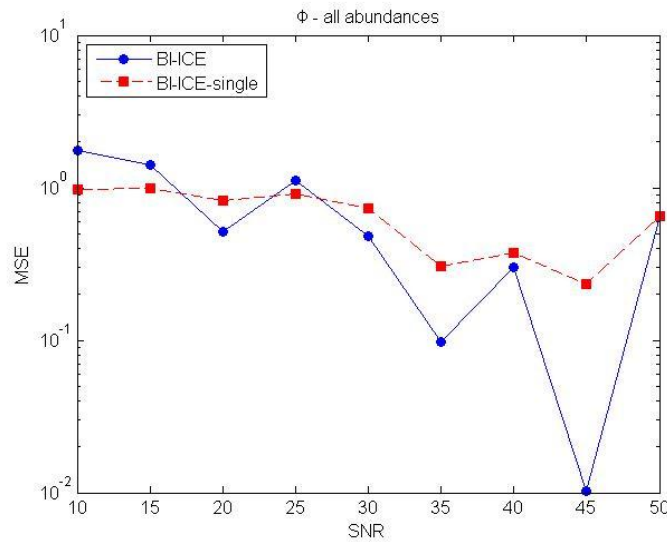


**Σχήμα 10:** Μέσο τετραγωνικό σφάλμα που προκύπτει λαμβάνοντας υπ' όψιν μόνο τα υλικά που συνθέτουν το εικονοστοιχείο, το οποίο παράγεται από τους BI-ICE και BI-ICE-single, για διαφόρους βαθμούς αραιότητας. Οι δύο αλγόριθμοι εφαρμόζονται σε προσομοιωμένα δεδομένα που προκύπτουν από το μητρώο  $\Phi_2$ , ενώ έχει προστεθεί λευκός θόρυβος (SNR = 20 dB).

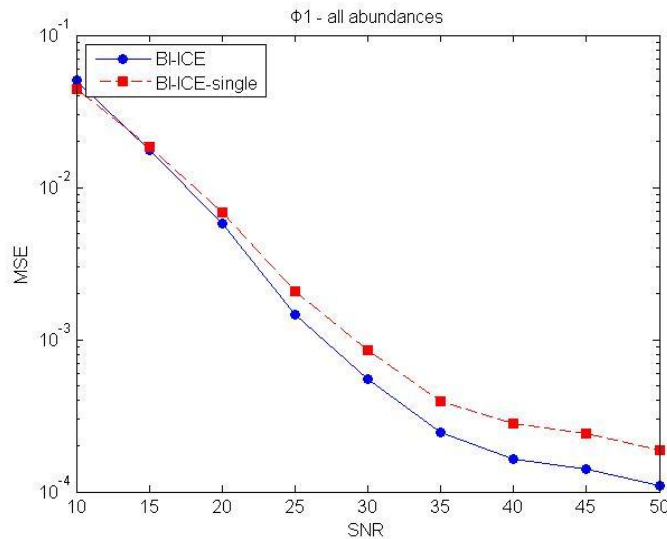
Από τη μελέτη των σχημάτων αυτών διαπιστώνουμε ότι τα συμπεράσματα που είχαμε εξάγει από τη μελέτη των σχημάτων που αφορούσαν το συνολικό διάνυσμα  $\vec{w}$ , εξακολουθούν να ισχύουν σε μεγάλο βαθμό και για τα Σχήματα 8, 9, 10. Καθώς αυξάνεται ο βαθμός αραιότητας επέρχεται μικρή αύξηση και στο σφάλμα εκτίμησης, και για τους δύο αλγόριθμους. Αυτό το φαινόμενο γίνεται πιο αισθητό στον BI-ICE, ενώ στην περίπτωση του BI-ICE-single, παρά τη μικρή αύξηση σφάλματος, που

παρατηρείται καθώς αυξάνεται ο βαθμός αραιότητας, δεν παρατηρούνται σημαντικές διακυμάνσεις στην τιμή του σφάλματος.

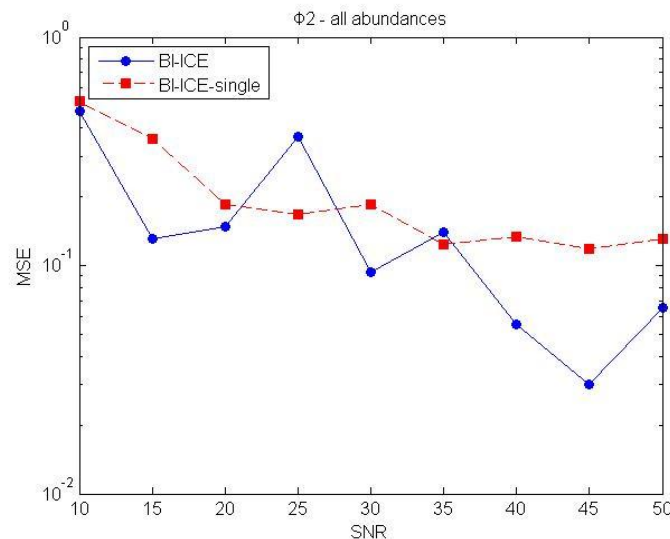
Στα Σχήματα 11-13 παρουσιάζονται οι τιμές του μέσου τετραγωνικού σφάλματος του διανύσματος ποσοστών συναρτήσει του λόγου σήματος-προς-θόρυβο, για τους δύο αλγορίθμους και για τα τρία μητρώα. Ο βαθμός αραιότητας για το διάνυσμα  $\vec{w}$  σε αυτά τα πειράματα είναι 8, και ο λόγος σήματος-προς-θόρυβο κυμαίνεται από 10 έως 50 dB.



**Σχήμα 11:** Μέσο τετραγωνικό σφάλμα που παράγεται από τους BI-ICE και BI-ICE-single σαν συνάρτηση του λόγου σήματος-προς-θόρυβο (dB), όταν (οι αλγόριθμοι) αυτοί εφαρμόζονται σε προσομοιωμένα δεδομένα που προκύπτουν από το μητρώο  $\Phi$ , ενώ έχει προστεθεί λευκός θόρυβος. Ο βαθμός αραιότητας  $\xi$  ισούται με 8.



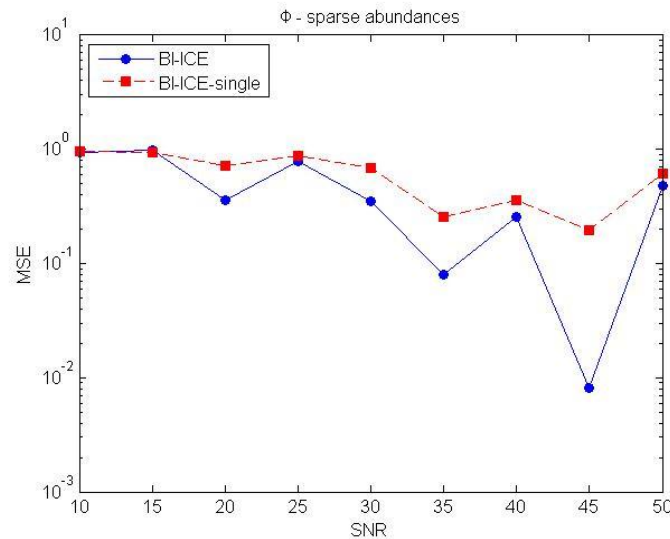
**Σχήμα 12:** Μέσο τετραγωνικό σφάλμα που παράγεται από τους BI-ICE και BI-ICE-single σαν συνάρτηση του λόγου σήματος-προς-θόρυβο (dB), όταν (οι αλγόριθμοι) αυτοί εφαρμόζονται σε προσομοιωμένα δεδομένα που προκύπτουν από το μητρώο  $\Phi_1$ , ενώ έχει προστεθεί λευκός θόρυβος. Ο βαθμός αραιότητας  $\xi$  ισούται με 8



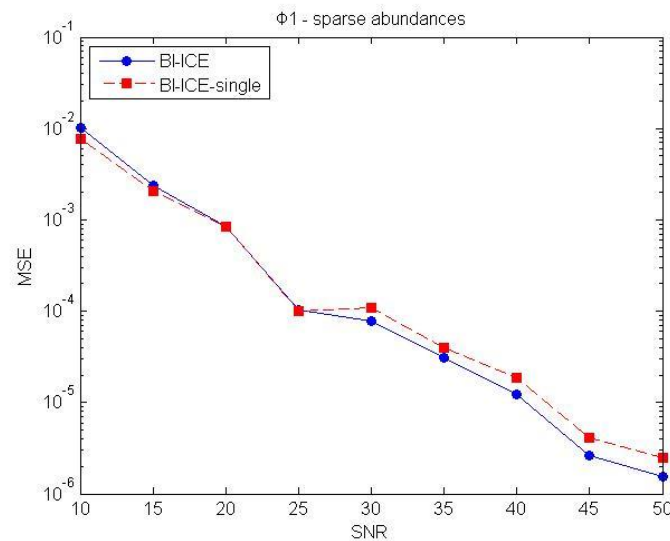
**Σχήμα 13:** Μέσο τετραγωνικό σφάλμα που παράγεται από τους BI-ICE και BI-ICE-single σαν συνάρτηση του λόγου σήματος-προς-θόρυβο (dB), όταν (οι αλγόριθμοι) αυτοί εφαρμόζονται σε προσομοιωμένα δεδομένα που προκύπτουν από το μητρώο  $\Phi_2$ , ενώ έχει προστεθεί λευκός θόρυβος. Ο βαθμός αραιότητας  $\xi$  ισούται με 8

Όπως αναμενόταν, οι τιμές του μέσου τετραγωνικού σφάλματος των αλγορίθμων μειώνονται, καθώς ο λόγος σήματος-προς-θόρυβο αυξάνεται. Κάτι τέτοιο γίνεται πιο ξεκάθαρο στην περίπτωση του μητρώου  $\Phi_1$ , που τα στοιχεία του ακολουθούν την κανονική κατανομή. Ακόμη, για άλλη μια φορά, παρατηρούμε ότι ο BI-ICE-single οδηγεί σε τιμές του μέσου τετραγωνικού σφάλματος, οι οποίες δεν παρουσιάζουν σημαντικές διακυμάνσεις, σε σχέση με τον BI-ICE.

Όπως και στην περίπτωση που εξετάσαμε το μέσο τετραγωνικό σφάλμα σε σχέση με το βαθμό αραιότητας, η μελέτη του σφάλματος, λαμβάνοντας υπ' όψιν μόνο τα στοιχεία του διανύσματος  $\vec{w}$  που αφορούν τα υλικά που βρίσκονται στο υπό εξέταση εικονοστοιχείο, θα μπορούσε να μας δώσει χρήσιμες πληροφορίες. Οι αντίστοιχες γραφικές παραστάσεις παρουσιάζονται στα Σχήματα 14-16.

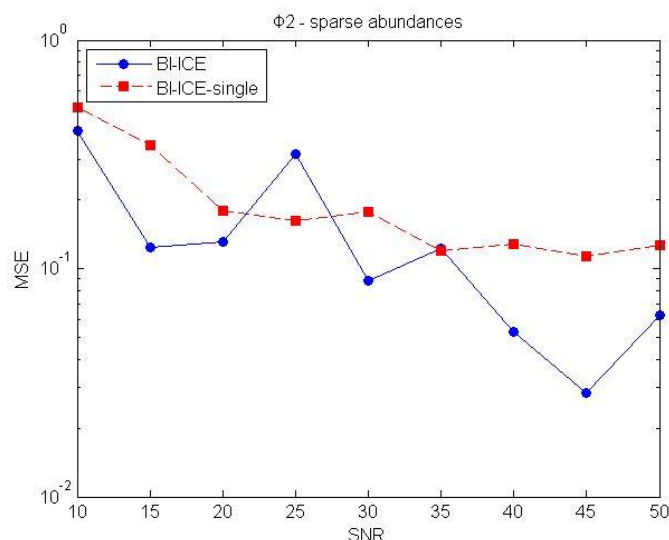


**Σχήμα 14:** Μέσο τετραγωνικό σφάλμα που παράγεται από τους δύο BI-ICE και BI-ICE-single σαν συνάρτηση του λόγου σήματος-προς-θόρυβο (dB), όταν (οι αλγόριθμοι) αυτοί εφαρμόζονται σε προσομοιωμένα δεδομένα που προκύπτουν από το μητρώο  $\Phi$ , ενώ έχει προστεθεί λευκός θόρυβος και λαμβάνονται υπ' όψιν μόνο τα υλικά που υπάρχουν στο εικονοστοιχείο. Ο βαθμός αραιότητας  $\xi$  ισούται με 8



**Σχήμα 15:** Μέσο τετραγωνικό σφάλμα που παράγεται από τους δύο BI-ICE και BI-ICE-single σαν συνάρτηση του λόγου σήματος-προς-θόρυβο (dB), όταν (οι αλγόριθμοι) αυτοί εφαρμόζονται σε προσομοιωμένα δεδομένα που προκύπτουν από το μητρώο  $\Phi_1$ , ενώ έχει προστεθεί λευκός θόρυβος και λαμβάνονται υπ' όψιν μόνο τα υλικά που υπάρχουν στο εικονοστοιχείο. Ο βαθμός αραιότητας  $\xi$  ισούται με 8





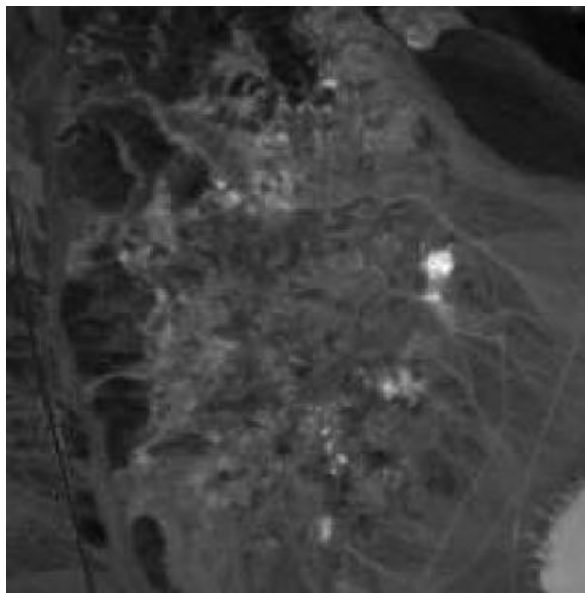
**Σχήμα 16:** Μέσο τετραγωνικό σφάλμα που παράγεται από τους BI-ICE και BI-ICE-single σαν συνάρτηση του λόγου σήματος-προς-θόρυβο (dB), όταν (οι αλγόριθμοι) αυτοί εφαρμόζονται σε προσομοιωμένα δεδομένα που προκύπτουν από το μητρώο  $\Phi_2$ , ενώ έχει προστεθεί λευκός θόρυβος και λαμβάνονται υπ' όψιν μόνο τα υλικά που υπάρχουν στο εικονοστοιχείο. Ο βαθμός αραιότητας  $\xi$  ισούται με 8

Παρατηρώντας τα Σχήματα 14-16, είναι φανερό ότι, και εδώ, το σφάλμα μειώνεται με την αύξηση του λόγου σήματος-προς-θόρυβο. Επιπλέον, κατά την αύξηση του λόγου σήματος-προς-θόρυβο παρατηρούμε ότι ο αλγόριθμος BI-ICE παράγει μικρότερα σφάλματα σε σχέση με τον BI-ICE-single, ιδίως στις περιπτώσεις των  $\Phi$  και  $\Phi_2$ . Αντιθέτως, όταν ο θόρυβος στο μοντέλο είναι μεγάλος φαίνεται ότι οι δύο αλγόριθμοι έχουν την ίδια απόδοση.

## 6.2 Αποτελέσματα προσομοίωσης σε πραγματικά δεδομένα

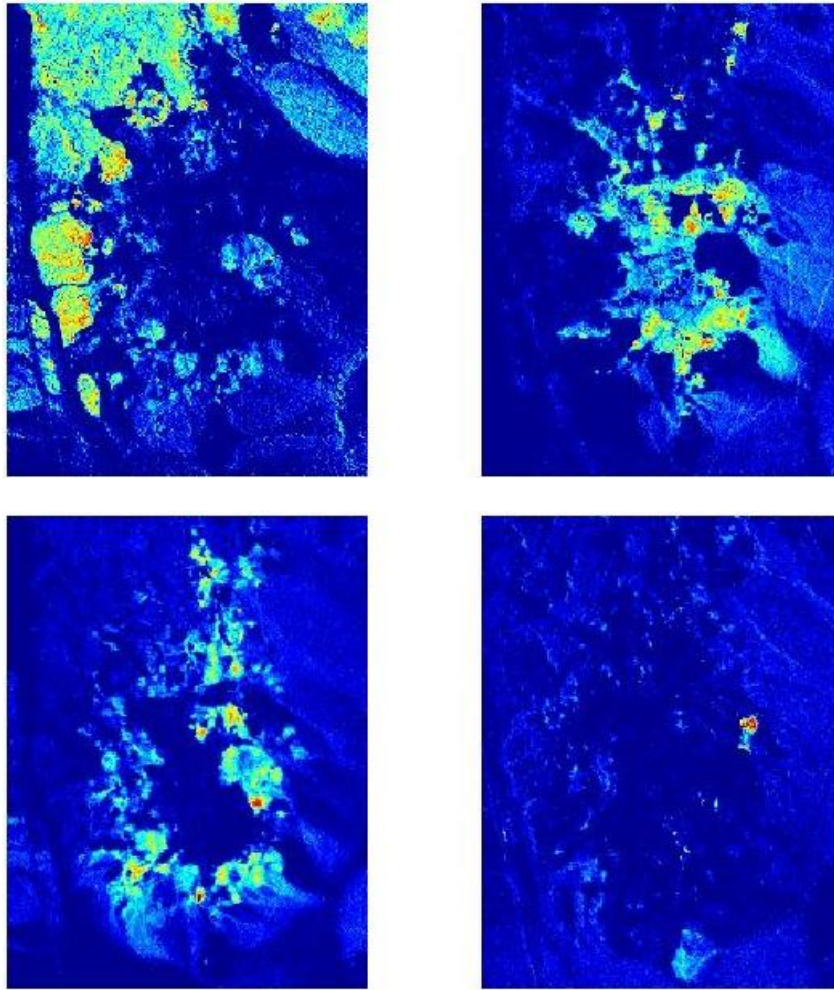
Αυτή η ενότητα περιλαμβάνει αποτελέσματα που προέκυψαν από την εφαρμογή της τεχνικής που προτείνεται στο Κεφάλαιο 5, η οποία λαμβάνει υπ' όψιν της τη χωρική συσχέτιση μεταξύ γειτονικών εικονοστοιχείων, πάνω σε πραγματικά δεδομένα (υπερφασματική εικόνα). Τα πραγματικά δεδομένα συλλέχθηκαν από την πτήση του εναέριου απεικονιστικού φασματόμετρου (imaging spectrometer) ορατών/υπέρυθρων ακτινοβολιών (AVIRIS) πάνω από το ορυχείο Cuprite της Νεβάδα, το 1997, [42]. Ο αισθητήρας του AVIRIS είναι ένα φασματόμετρο απεικόνισης 224 καναλιών, με περίπου 10 nm φασματική ανάλυση, καλύπτοντας μήκη κύματος εύρους από 0,4 έως 2,5  $\mu\text{m}$ . Η χωρική του ανάλυση είναι 20 m. Το συγκεκριμένο σύνολο δεδομένων έχει χρησιμοποιηθεί ευρέως για πειράματα τηλεπισκόπησης [10], [43]-[45]. Οι φασματικές

ζώνες 1-2, 104-113, 148-167, και 221-224 αφαιρέθηκαν λόγω χαμηλού SNR και απορρόφησης υδρατμών. Ως εκ τούτου, λαμβάνονται υπ' όψιν συνολικά 188 φασματικές ζώνες σε αυτό το πείραμα. Η υπό-εικόνα της 150<sup>ης</sup> ζώνης συχνοτήτων (band), η οποία περιλαμβάνει 200 κατακόρυφες γραμμές με 200 δείγματα ανά γραμμή (200 x 200) φαίνεται στην Εικόνα 2.

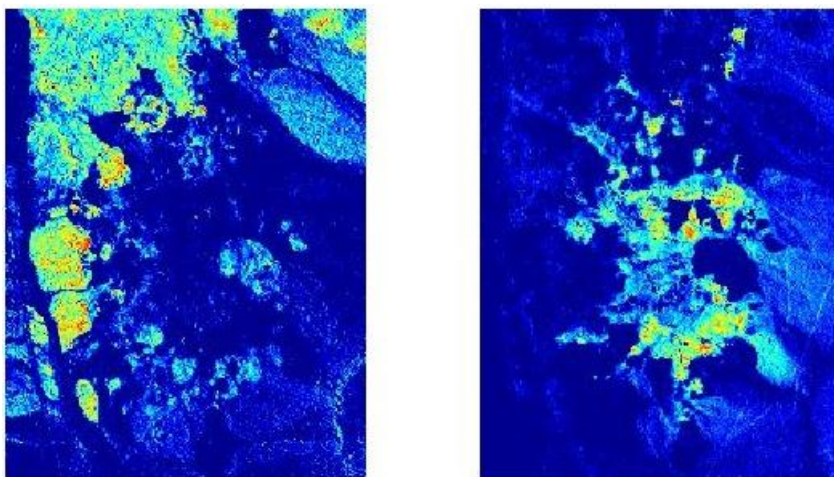


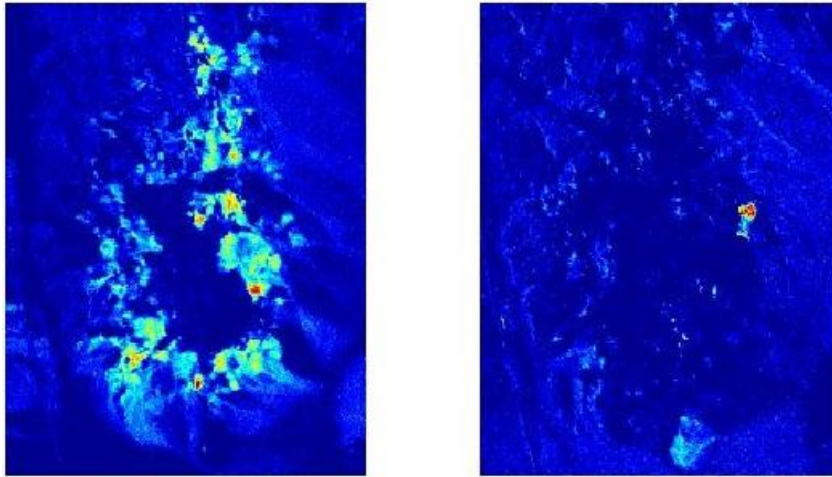
**Εικόνα 2: Υπό-εικόνα της 150<sup>ης</sup> ζώνης συχνοτήτων του υπερφασματικού συνόλου δεδομένων Cuprite AVIRIS**

Ο αλγόριθμος VCA χρησιμοποιήθηκε για την εξαγωγή 14 υλικών που εμφανίζονται στην εικόνα, όπως στο [10]. Χρησιμοποιώντας τις φασματικές υπογραφές των υλικών αυτών, οι δύο αλγόριθμοι BI-ICE και BI-ICE-single, δοκιμάζονται τόσο στο πλαίσιο χωρικής όσο και στο πλαίσιο της μη χωρικής συσχέτισης, για την εκτίμηση των αντίστοιχων ποσοστών σε κάθε εικονοστοιχείο. Η διαδικασία φασματικού διαχωρισμού παράγει μια εικόνα για κάθε υλικό, που απεικονίζει την εκτιμώμενη ποσοστιαία αναλογία του υλικού σε κάθε εικονοστοιχείο. Όσο πιο σκούρο είναι το χρώμα του εικονοστοιχείου, τόσο μικρότερη είναι η συμβολή του υλικού στο εν λόγω εικονοστοιχείο. Αντίθετα, ένα εικονοστοιχείο ανοικτού χρώματος υποδηλώνει ότι το ποσοστό συμμετοχής του υλικού στο συγκεκριμένο εικονοστοιχείο είναι υψηλό. Στις Εικόνες 3, 4, 5, 6 παρουσιάζονται οι ποσοστιαίες αναλογίες τεσσάρων υλικών, που έχουν εκτιμηθεί με τη χρήση των αλγορίθμων (α) BI-ICE χωρίς χωρική συσχέτιση, (β) BI-ICE με χωρική συσχέτιση, (γ) BI-ICE-single χωρίς χωρική συσχέτιση και (δ) BI-ICE-single με χωρική συσχέτιση, αντίστοιχα.

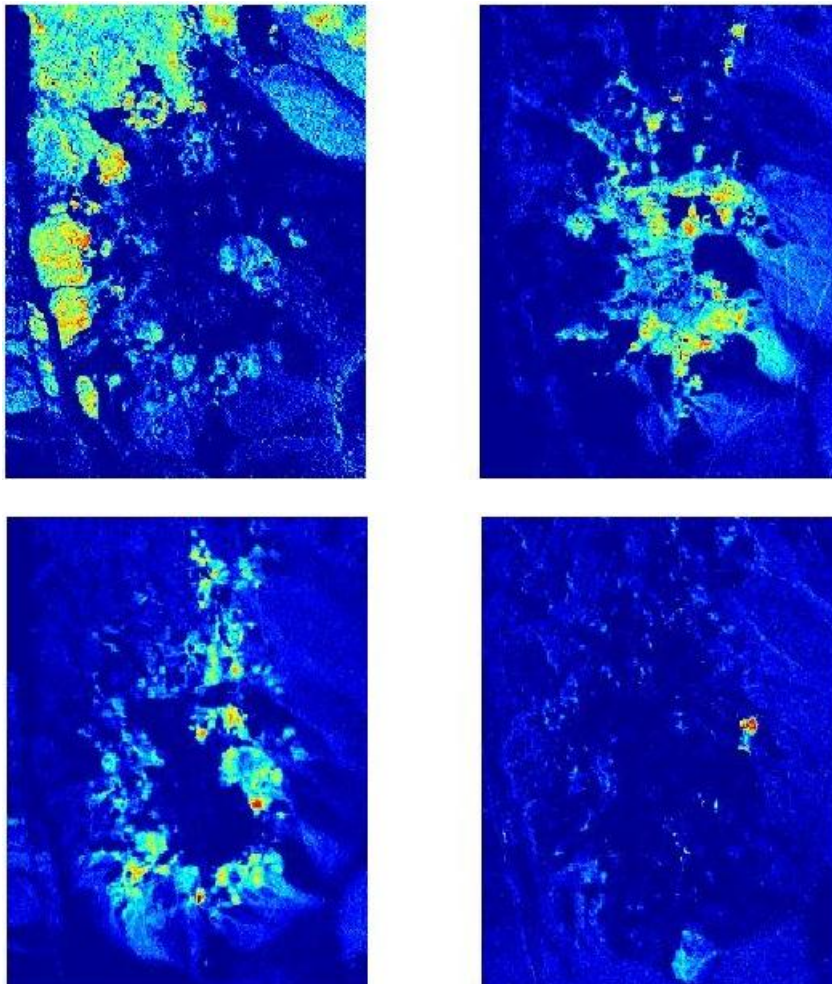


**Εικόνα 3: Εκτιμώμενες τιμές ποσοστών για τέσσερα υλικά, χρησιμοποιώντας τον αλγόριθμο BI-ICE χωρίς χωρική συσχέτιση**



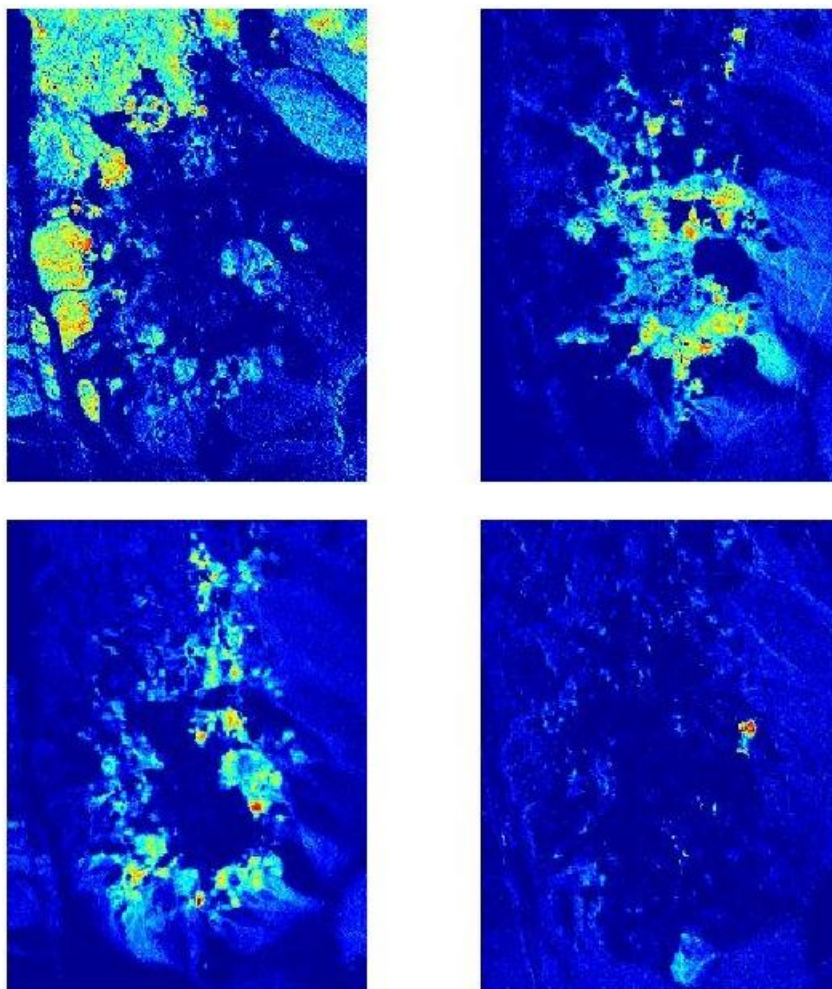


**Εικόνα 4: Εκτιμώμενες τιμές ποσοστών για τέσσερα υλικά, χρησιμοποιώντας τον αλγόριθμο BI-ICE με χωρική συσχέτιση**



**Εικόνα 5: Εκτιμώμενες τιμές ποσοστών για τέσσερα υλικά, χρησιμοποιώντας τον αλγόριθμο BI-ICE-single χωρίς χωρική συσχέτιση**



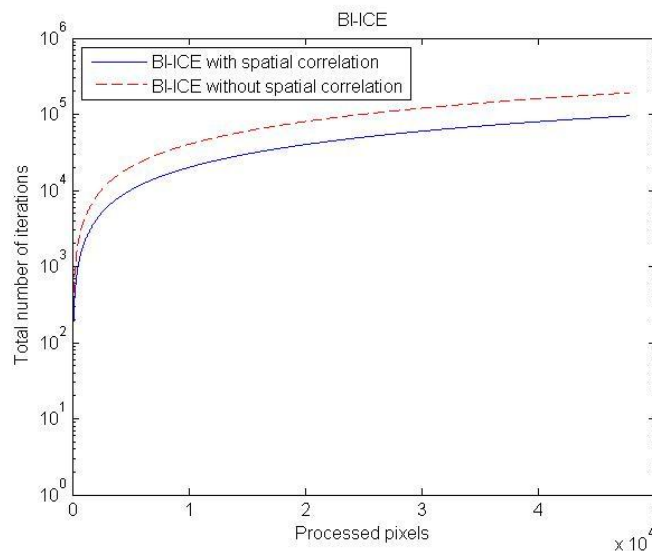


**Εικόνα 6: Εκτιμώμενες τιμές ποσοστών για τέσσερα υλικά, χρησιμοποιώντας τον αλγόριθμο BI-ICE-single με χωρική συσχέτιση**

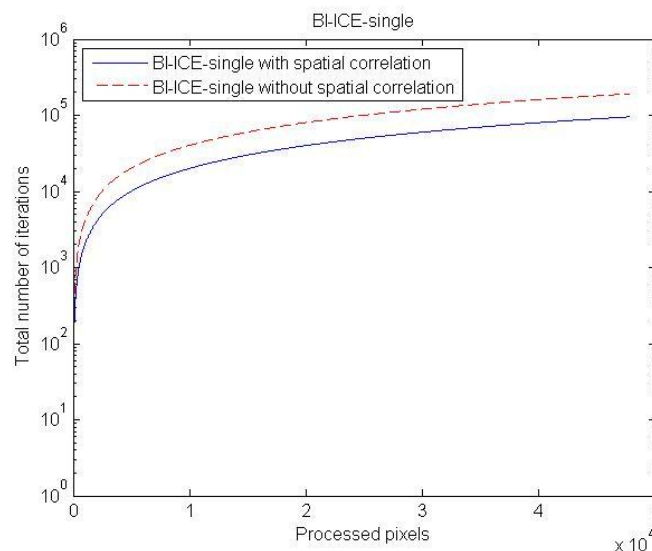
Από τις εικόνες αυτές, φαίνεται ότι οι αλγόριθμοι BI-ICE και BI-ICE-single, τόσο με όσο και χωρίς χωρική συσχέτιση, οδηγούν σε αποτελέσματα που παρουσιάζουν μεγάλο βαθμό ομοιότητας, και είναι σε πλήρη συμφωνία με προηγούμενα αντίστοιχα αποτελέσματα που έχουν δημοσιευθεί στα [10], [45], όπως και με τα συμπεράσματα που προκύπτουν από το εδάφιο 6.1.

Επιπλέον, για την αξιολόγηση της απόδοσης των αλγορίθμων, όσον αφορά τις απαιτήσεις σε υπολογιστική ισχύ, εξετάζεται ο συνολικός αριθμός επαναλήψεων, που χρειάστηκαν για την επίτευξη σύγκλισης, για κάθε ένα από τα εικονοστοιχεία της εικόνας. Στα Σχήματα 17 και 18, φαίνεται σε ημί-λογαριθμική κλίμακα πώς ο συνολικός αριθμός επαναλήψεων αυξάνεται κατά τη διάρκεια της εκτέλεσης. Πιο συγκεκριμένα, στο Σχήμα 17 απεικονίζεται με κόκκινο χρώμα και διακεκομμένη γραμμή η πορεία του συνολικού αριθμού των επαναλήψεων για τον αλγόριθμο BI-ICE χωρίς χωρική συσχέτιση και με μπλε χρώμα και συνεχή γραμμή ο αντίστοιχος αριθμός για τον

αλγόριθμο BI-ICE με χωρική συσχέτιση, καθώς σαρώνονται τα εικονοστοιχεία της εικόνας. Επιπλέον, στο Σχήμα 18 φαίνεται η αντίστοιχη γραφική απεικόνιση για τον αλγόριθμο BI-ICE-single, τόσο με όσο και χωρίς χωρική συσχέτιση (μπλε χρώμα και συνεχή γραμμή, κόκκινο χρώμα και διακεκομμένη γραμμή, αντίστοιχα).



**Σχήμα 17:** Γραφική απεικόνιση του συνολικού αριθμού επαναλήψεων κατά τη διάρκεια του πειράματος, για τον αλγόριθμο BI-ICE με και χωρίς χωρική συσχέτιση (μπλε συνεχής γραμμή και κόκκινη διακεκομμένη γραμμή, αντίστοιχα)



**Σχήμα 18:** Γραφική απεικόνιση του συνολικού αριθμού επαναλήψεων κατά τη διάρκεια του πειράματος, για τον αλγόριθμο BI-ICE-single με και χωρίς χωρική συσχέτιση (μπλε συνεχής γραμμή και κόκκινη διακεκομμένη γραμμή, αντίστοιχα)

Από την παρατήρηση των γραφημάτων προκύπτει εύκολα ότι η τεχνική της συσχέτισης επιτυγχάνει σημαντική μείωση του αριθμού των επαναλήψεων που απαιτούνται για σύγκλιση, για κάθε εικονοστοιχείο. Αυτό αιτιολογείται λαμβάνοντας υπ' όψιν ότι γειτονικά εικονοστοιχεία είναι πολύ πιθανό να περιέχουν κατά ένα ποσοστό κοινά σε αριθμό ή/και σε αναλογίες υλικά (να έχουν δηλαδή παρόμοια μορφή). Έτσι, η αξιοποίηση της πληροφορίας, που ήδη υπάρχει από προηγούμενα γειτονικά εικονοστοιχεία, έχει ως αποτέλεσμα τον περιορισμό των επαναλήψεων σε σχέση με την περίπτωση όπου δε λαμβάνεται υπ' όψιν η χωρική συσχέτιση, χωρίς όμως να προκύπτει υποβάθμιση της ποιότητας των αποτελεσμάτων, όπως φαίνεται από τη σύγκριση των Εικόνων 3, 4, 5, 6.

**Πίνακας 3: Μέσος αριθμός επαναλήψεων, συνολικός χρόνος εκτέλεσης και μέσος χρόνος εκτέλεσης ανά εικονοστοιχείο, για τους αλγορίθμους BI-ICE και BI-ICE-single, με και χωρίς χωρική συσχέτιση**

	Μέσος αριθμός επαναλήψεων ανά εικονοστοιχείο	Συνολικός χρόνος εκτέλεσης αλγορίθμου (δευτερόλεπτα) <sup>7</sup>	Μέσος χρόνος εκτέλεσης ανά εικονοστοιχείο (δευτερόλεπτα)
BI-ICE χωρίς χωρική συσχέτιση	4,0057	231,76	0,004854
BI-ICE με χωρική συσχέτιση	2	79,62	0,001667
BI-ICE-single χωρίς χωρική συσχέτιση	4,0057	226,40	0,004741
BI-ICE-single με χωρική συσχέτιση	2	77,57	0,001624

<sup>7</sup> Οι τιμές είναι ενδεικτικές. Οι πειραματικές δοκιμές έγιναν σε σύστημα MS Windows 7 – 64-bit, CPU Intel Core i7 Q740 1,73 GHz, RAM 6 GB και έκδοση MatLab R2010b.

Στον Πίνακα 3 παρουσιάζονται (α) ο μέσος αριθμός επαναλήψεων που απαιτήθηκαν για την εκτέλεση του πειράματος για κάθε έναν από τους αλγόριθμους, (β) ο απαιτούμενος χρόνος για την ολοκλήρωση της διαδικασίας και (γ) ο μέσος χρόνος που απαιτείται για την επίτευξη σύγκλισης για ένα εικονοστοιχείο.



## 7. ΣΥΜΠΕΡΑΣΜΑΤΑ

Στα πλαίσια της παρούσας εργασίας εξετάστηκε το πρόβλημα του ημί-επιβλεπόμενου υπερφασματικού διαχωρισμού, το οποίο εκφράστηκε με τη χρήση ενός ιεραρχικού μοντέλου κατά Bayes, όπου οι περιορισμοί και οι ιδιότητες των παραμέτρων του ενσωματώνονται επιλέγοντας κατάλληλες εκ των προτέρων κατανομές πιθανότητας για αυτές, ώστε να εξασφαλίζεται τόσο η αραιότητα του διανύσματος των ποσοστών συμμετοχής των υλικών στη σύνθεση ενός εικονοστοιχείου, όσο και η μη αρνητικότητα του. Στη συνέχεια, αναπτύχθηκε ένας νέος επαναληπτικός αλγόριθμος συμπερασμού κατά Bayes, για την εκτίμηση των παραμέτρων του μοντέλου, με όνομα BI-ICE-single, ο οποίος παρουσιάζει μεγάλο βαθμού συγγένεια με τον αλγόριθμο BI-ICE.

Αξίζει να σημειωθεί ότι τόσο ο BI-ICE-single, όσο και ο BI-ICE, μπορούν να θεωρηθούν ως μία ντετερμινιστική προσέγγιση του δειγματολήπτη Gibbs, [23]. Σε αντίθεση με ότι συμβαίνει στη δειγματοληψία κατά Gibbs, όπου παράγονται τυχαία δείγματα με επαναληπτικό τρόπο από τις υπό συνθήκη εκ των υστέρων κατανομές των παραμέτρων του μοντέλου, στους αλγορίθμους αυτούς, χρησιμοποιούνται οι εκ των υστέρων υπό συνθήκη μέσες τιμές των αντίστοιχων υπό συνθήκη εκ των υστέρων κατανομών των παραμέτρων.

Ο BI-ICE-single δεν παρουσιάζει την ίδια ακρίβεια στον υπολογισμό των ποσοστών των υλικών με τον BI-ICE, όταν οι φασματικές υπογραφές αυτών (των υλικών) παρουσιάζουν υψηλό βαθμό συσχέτισης. Αντίθετα, όταν ο βαθμός συσχέτισης των παραπάνω φασματικών υπογραφών δεν είναι υψηλός, οι δύο αλγόριθμοι παρουσιάζουν παρόμοια συμπεριφορά.

Επιπλέον, προτάθηκε μια νέα απλή τεχνική για την εκτίμηση των ποσοστών συμμετοχής του κάθε υλικού στα εικονοστοιχεία μιας υπερφασματικής εικόνας, η οποία λαμβάνει υπ' όψιν της την πιθανή χωρική συσχέτιση μεταξύ γειτονικών εικονοστοιχείων. Η τεχνική αυτή παρουσιάζει την ίδια ακρίβεια στις εκτιμήσεις των διανυσμάτων ποσοστών, σε σχέση με την περίπτωση όπου δεν λαμβάνεται υπ' όψιν η πιθανή χωρική συσχέτιση μεταξύ γειτονικών εικονοστοιχείων. Ταυτόχρονα, οδηγεί σε σημαντική εξοικονόμηση υπολογιστικής ισχύος.

## ΠΙΝΑΚΑΣ ΟΡΟΛΟΓΙΑΣ

Ξενόγλωσσος όρος	Ελληνικός Όρος
Abundance	Ποσοστό συμμετοχής υλικού στη δημιουργία ενός εικονοστοιχείου
Abundance vector	Διάνυσμα ποσοστών συμμετοχής υλικών στη δημιουργία ενός εικονοστοιχείου
Additive white Gaussian noise	Προσθετικός λευκός Gaussian θόρυβος
Additivity constraint	Προσθετικός περιορισμός
Band	Ζώνη συχνοτήτων
Bayesian inference	Συμπερασμός κατά Bayes
Compressive sensing techniques	Τεχνικές αναπαράστασης αραιών σημάτων
Condition number	Αριθμός συνθήκης
Conjugacy	Συζυγία
Electro-optical remote sensing	Ηλεκτρο-οπτική τηλεπισκόπηση
Endmember	Υλικό (φασματική υπογραφή) που συμμετέχει στη δημιουργία (της φασματικής υπογραφής) ενός εικονοστοιχείου.
Extraction technique	Τεχνική εξαγωγής χαρακτηριστικών
Fully constrained least squares method	Η μέθοδος των ελαχίστων τετραγώνων που λαμβάνει υπ' όψιν όλους τους περιορισμούς
Gamma distribution	Γάμμα κατανομή
Gaussian distribution	Κανονική κατανομή
Gaussian scale mixture	Σταθμισμένη μίξη κανονικών κατανομών
Generalized constraint	Γενικευμένος περιορισμός
Generalized inverse distribution	Γενικευμένη αντίστροφη κατανομή
Homogeneous mixture	Ομοιογενής μίξη
Hyperparameter	Υπέρ-παράμετρος
Hyperspectral imagery	Υπερφασματική απεικόνιση
Hyperspectral imaging sensors	Υπερφασματικοί αισθητήρες απεικόνισης
Imaging spectrometer	Απεικονιστικό φασματόμετρο

Imaging spectroscopy	Απεικονιστική φασματοσκοπία
Independent identically distributed	Ανεξάρτητα ομοιόμορφα κατανεμημένα
Laplace distribution	Λαπλασιανή κατανομή
Linear mixture model	Γραμμικό μοντέλο μίξης
Marginalization	Περιθωριοποίηση
Maximum a posteriori estimation	Μέγιστη εκ των υστέρων εκτίμηση
Mean squared error	Μέσο τετραγωνικό σφάλμα
Mutual coherence	Αμοιβαία συνάφεια
Nonnegative constraint	Περιορισμός της μη αρνητικότητας
Pixel	Εικονοστοιχείο
Posterior conditional expectation	Εκ των υστέρων υπό συνθήκη μέσες τιμές
Prior probability distribution ή prior	Εκ των προτέρων κατανομές πιθανότητας
Probability density function	Συνάρτηση πυκνότητας πιθανότητας
Remote scene	Απομακρυσμένη σκηνή
Scale factor	Συντελεστής κλίμακας
Semi-supervised	Ημί-επιβλεπόμενος
Signal-to-noise ratio	Λόγος σήματος-προς-θόρυβο
Signature matrix	Πίνακας υπογραφών
Simplex	Μονόπλοκο
Sparsity	Αραιότητα
Spatial resolution	Χωρική ανάλυση
Spectral bands	Φασματικές ζώνες
Spectral dimension	Φασματική διάσταση
Spectral signature	Φασματική υπογραφή
Spectral unmixing	Φασματικός διαχωρισμός
Spectroscopy	Φασματοσκοπία
Supervised	Επιβλεπόμενος
Terrain classifications	Ταξινόμηση τύπου εδάφους
Truncated Gaussian distribution	Κολοβωμένη κανονική κατανομή
Unimodality	Ιδιότητα μοναδικής κορυφής
Unsupervised	Μη επιβλεπόμενος
Waveband	Ζώνη συχνοτήτων
Wavelengths	Μήκη κύματος
Wide-area	Ευρεία ζώνη

## ΣΥΝΤΜΗΣΕΙΣ – ΑΡΚΤΙΚΟΛΕΞΑ – ΑΚΡΩΝΥΜΙΑ

AVIRIS	Airborne Visible/Infrared Imaging Spectrometer
AWGN	Additive White Gaussian Noise
BI-ICE	Bayesian Inference Iterative Conditional Expectations
GSM	Gaussian Scale Mixture
i.i.d.	Independent Identically Distributed
Lasso	Least Absolute Shrinkage and Selection Operator
LMM	Linear Mixture Model
MAP	Maximum A Posteriori estimation
MSE	Mean Squared Error
pdf	Probability Density Function
PPI	Pixel Purity Index
SNR ή S/N	Signal-to-Noise Ratio
SU	Spectral Unmixing
USGS	United States Geological Survey
VCA	Vertex Component Analysis
EAA	Εθνικό Αστεροσκοπείο Αθηνών
ΕΚΠΑ	Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών

## ΠΑΡΑΡΤΗΜΑ – ΚΩΔΙΚΑΣ MATLAB

Στα πλαίσια της παρούσας εργασίας αναπτύχθηκε κώδικας, απαραίτητος για την υλοποίηση και πειραματική μελέτη των αναφερόμενων αλγορίθμων και τεχνικών. Ακολουθούν τα αρχεία κώδικα MatLab, όπως αυτά είναι οργανωμένα και στην ηλεκτρονική μορφή της εργασίας.

### Συναρτήσεις υλοποίησης αλγορίθμων:

#### BiICE.m:

```
function [ w_BiICE, ws, vita, gm, lamda ] = BiICE ( Phi, y, MaxIter, aVita,
bVita, aLamda, bLamda )
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This code implements the fast Laplace algorithm from the following paper:
%
% [1] K. E. Themelis, A. A. Rontogiannis, K. D. Koutroumbas. "A Novel
Hierarchical Bayesian Approach for Sparse Semi-Supervised Hyperspectral
Unmixing,"
% submitted for publication, IEEE Transactions on Signal Processing, December
2010.
%
% Function:
%
% [ w_BiICE, ws, vita, gm, lamda ] = BiICE ( Phi, y, MaxIter, aVita, bVita,
aLamda, bLamda )
%
% Inputs:
% PHI:      measurement matrix
% y:        vector of measurements
% MaxIter:  number of maximum iterations
% aVita:    the shape parameter of the gamma distribution of vita
% bVita:    the inverse of the scale parameter of the gamma distribution
%           of vita
% aLamda:   the shape parameter of the gamma distribution of lamda
% bLamda:   the inverse of the scale parameter of the gamma distribution
%           of lamda
%
% Outputs:
% w_BiICE:  final estimation vector
% ws:       conditional means for w
% vita:     conditional means for vita
% gm:       conditional means for gamma
% lamda:    conditional means for lamda
%
% (c) 2010 K. E. Themelis, A. A. Rontogiannis, K. D. Koutroumbas
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

[M,N] = size(Phi);

% compute the covariance matrix
Phi_cov = Phi.' * Phi;

% initialization
ws = ones (N,MaxIter); % conditional means for w
gm = ones (N,MaxIter); % conditional means for gamma
L = eye (N);
```

```

lamda = ones(N,MaxIter); % conditional means for lamda
vita = ones(MaxIter,1); % conditional means for vita

for q = 2 : MaxIter

    % sample the gaussian distribution of w
    Sw = vita(q-1)^-1 * inv( Phi_cov + L);
    Sw_inv = vita(q-1) * ( Phi_cov + L);
    mw = vita(q-1) * Sw * Phi.' * y;
    ws_prev = mw;

    for t = 1 : N

        idx1 = cell(1);
        idx1{1} = mod((0:N-1)+1, N)+1;
        ws_prev = ws_prev(idx1{:});
        mw = mw(idx1{:});

        idx2 = {idx1{1} idx1{1}};
        Sw = Sw(idx2{:});
        Sw_inv = Sw_inv(idx2{:});
        S1 = Sw(1:end-1,end);

        tmp2 = [S1.' 0] * Sw_inv ;
        qalpha = tmp2 * [S1; 0];
        tmp3 = Sw(end,end) / (qalpha + Sw(end,end));
        st = Sw(end,end) - qalpha * tmp3;
        mt = mw(end) + tmp2 * [ws_prev(1:end-1)-mw(1:end-1); 0] * tmp3;

        tmp1 = 1 - .5 * erfc(mt/real(sqrt(2))/sqrt(st));
        if abs(tmp1) < 1e-9
            ws_prev(end) = 1e-121;
        else
            ws_prev(end) = mt + (tmp1)^-1 * (2*pi)^-(1/2) * sqrt(st) * exp(-
.5*mt^2/st);
        end
    end
    ws(:,q) = ws_prev;

    % compute the mean of gamma distribution of vita
    vita(q) = (M/2 + aVita + N/2) / (1/2 * norm(y-Phi*ws(:,q),2)^2 + bVita +
1/2 * ws(:,q).' * L * ws(:,q));

    % compute the mean of distribution of gamma
    gm(:,q) = sqrt(2*lamda(:,q-1)./pi).*exp(sqrt(lamda(:,q-
1).*vita(q)).*(ws(:,q))).*(vita(q)*ws(:,q).^2./lamda(:,q-
1)).^(3/4).*besselk(3/2,sqrt(lamda(:,q-1).*vita(q)).*ws(:,q)));
    L = diag(1./gm(:,q));

    % compute the mean of gamma distribution of lamda
    lamda(:,q) = (1 + aLamda)./(1/2 * gm(:,q) + bLamda);

end
w_BiICE = ws(:,end);

```

**BiICE\_single\_lamda.m:**

```

function [ w_BiICE, ws, vita, gm, lamda ] = BiICE_single_lamda ( Phi, y,
MaxIter, aVita, bVita, aLamda, bLamda )
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

```

```

% This code implements the fast Laplace algorithm from the following paper:
%[1] K. E. Themelis, A. A. Rontogiannis, K. D. Koutroumbas. "A Novel
%Hierarchical Bayesian Approach for Sparse Semi-Supervised Hyperspectral
%Unmixing," submitted for publication, IEEE Transactions on Signal
%Processing, December 2010.
%
% Function:
%
% [ w_BiICE, ws, vita, gm, lamda ] = BiICE_single_lamda ( Phi, y, MaxIter,
%aVita, bVita, aLamda, bLamda )
%
% Inputs:
% PHI:      measurement matrix
% y:        vector of measurements
% MaxIter:  number of maximum iterations
% aVita:    the shape parameter of the gamma distribution of vita
% bVita:    the inverse of the scale parameter of the gamma distribution
%           of vita
% aLamda:   the shape parameter of the gamma distribution of lamda
% bLamda:   the inverse of the scale parameter of the gamma distribution
%           of lamda
%
% Outputs:
% w_BiICE:  final estimation vector
% ws:       conditional means for w
% vita:     conditional means for vita
% gm:       conditional means for gamma
% lamda:    conditional means for lamda
%
% (c) 2010 K. E. Themelis, A. A. Rontogiannis, K. D. Koutroumbas
% Modified: (c) June 2012 Angelopoulos Vasileios, Kafouris Pavlos
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

[M,N] = size(Phi);

% Compute the covariance matrix
Phi_cov = Phi.' * Phi;

% Initialization
ws = ones(N,MaxIter); % Conditional means for w
gm = ones(N,MaxIter); % Conditional means for gamma
L = eye(N);
lamda = ones(1,MaxIter); % Conditional means for lambda
vita = ones(MaxIter,1); % Conditional means for vita

for q = 2 : MaxIter

    % Sample the gaussian distribution of w
    Sw = vita(q-1)^-1 * inv( Phi_cov + L);
    Sw_inv = vita(q-1) * ( Phi_cov + L);
    mw = vita(q-1) * Sw * Phi.' * y;
    ws_prev = mw;

    for t = 1 : N

        idx1 = cell(1);
        idx1{1} = mod((0:N-1)+1, N)+1;
        ws_prev = ws_prev(idx1{:});
        mw = mw(idx1{:});

        idx2 = {idx1{1} idx1{1}};
        Sw = Sw(idx2{:});
        Sw_inv = Sw_inv(idx2{:});
    end
end

```

```

S1 = Sw(1:end-1,end);

tmp2 = [S1.' 0] * Sw_inv ;
qalpha = tmp2 * [S1; 0];
tmp3 = Sw(end,end) / (qalpha + Sw(end,end));
st = Sw(end,end) - qalpha * tmp3;
mt = mw(end) + tmp2 * [ws_prev(1:end-1)-mw(1:end-1); 0] * tmp3;

tmp1 = 1 - .5 * erfc(mt/realsqrt(2)/sqrt(st));
if abs(tmp1) < 1e-9
    ws_prev(end) = 1e-121;
else
    ws_prev(end) = mt + (tmp1)^-1 * (2*pi)^-(1/2) * sqrt(st) * exp(-
.5*mt^2/st);
end
end
ws(:,q) = ws_prev;

% Compute the mean of gamma distribution of vita
vita(q) = (M/2 + aVita + N/2) / (1/2 * norm(y-Phi*ws(:,q),2)^2 + bVita +
1/2 * ws(:,q).' * L * ws(:,q));

% Compute the mean of distribution of gamma
gm(:,q) = sqrt(2*lamda(1,q-1)./pi).*exp(sqrt(lamda(1,q-
1).*vita(q)).*(ws(:,q))).*(vita(q)*ws(:,q).^2./lamda(1,q-
1)).^(3/4).*besselk(3/2,sqrt(lamda(1,q-1).*vita(q)).*ws(:,q)));
L = diag(1./gm(:,q));

% Compute lamda
lamda(1,q) = (N + aLamda)./(sum(1/2 * gm(:,q)) + bLamda);

end
w_BiICE = ws(:,end);

```

### **BiICE pixel correlation.m:**

```

function [ w_BiICE, vita_fin, gm_fin, lamda_fin, q ] =
BiICE_pixel_correlation ( Phi, y, MaxIter, w_pr, vita_pr, gm_pr, lamda_pr,
aVita, bVita, aLamda, bLamda )
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This code implements the fast Laplace algorithm from the following paper:
%
% [1] K. E. Themelis, A. A. Rontogiannis, K. D. Koutroumbas. "A Novel
%Hierarchical Bayesian Approach for Sparse Semi-Supervised Hyperspectral
%Unmixing," submitted for publication, IEEE Transactions on Signal
%Processing, December 2010.
%
% Function:
%
% [ w_BiICE, vita_fin, gm_fin, lamda_fin, q ] = BiICE_pixel_correlation( Phi,
y, MaxIter, w_pr, vita_pr, gm_pr, lamda_pr, aVita, bVita, aLamda, bLamda )
% Inputs:
% PHI:      measurement matrix
% y:       vector of measurements
% MaxIter:  number of maximum iterations
% w_pr:    the final estimation of the correlated pixel
% vita_pr:  the final vita of the correlated pixel
% gm_pr:    the final gm of the correlated pixel
% lamda_pr: the final lamda of the correlated pixel
% aVita:    the shape parameter of the gamma distribution of vita
% bVita:    the inverse of the scale parameter of the gamma distribution

```



```

%           of vita
%   aLamda:  the shape parameter of the gamma distribution of lamda
%   bLamda:  the inverse of the scale parameter of the gamma distribution
%           of lamda
% Outputs:
%   w_BiICE: final estimation vector
%   vita_fin: conditional means for vita in the last step
%   gm_fin:  conditional means for gamma in the last step
%   lamda_fin: conditional means for lamda in the last step
%   q:      total number of iterations that were necessary for
%           convergence
%
% (c) 2010 K. E. Themelis, A. A. Rontogiannis, K. D. Koutroumbas
% Modified: (c) June 2012 Angelopoulos Vasileios, Kafouris Pavlos
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

[M,N] = size(Phi);

% Compute the covariance matrix
Phi_cov = Phi.' * Phi;

% Initialization
ws = ones(N,MaxIter); % Conditional means for w
ws(:,1) = w_pr;
gm = ones(N,MaxIter); % Conditional means for gamma
gm(:,1) = gm_pr;
L = diag(1./gm(:,1));
lamda = ones(N,MaxIter); % Conditional means for lambda
lamda(:,1) = lamda_pr;
vita = ones(MaxIter,1); % Conditional means for vita
vita(1,1) = vita_pr;

for q = 2 : MaxIter

    % Sample the gaussian distribution of w
    Sw = vita(q-1)^-1 * inv( Phi_cov + L);
    Sw_inv = vita(q-1) * ( Phi_cov + L);
    mw = vita(q-1) * Sw * Phi.' * y;
    ws_prev = mw;

    for t = 1 : N

        idx1 = cell(1);
        idx1{1} = mod((0:N-1)+1, N)+1;
        ws_prev = ws_prev(idx1{:});
        mw = mw(idx1{:});

        idx2 = {idx1{1} idx1{1}};
        Sw = Sw(idx2{:});
        Sw_inv = Sw_inv(idx2{:});
        S1 = Sw(1:end-1,end);

        tmp2 = [S1.' 0] * Sw_inv ;
        qalpha = tmp2 * [S1; 0];
        tmp3 = Sw(end,end) / (qalpha + Sw(end,end));
        st = Sw(end,end) - qalpha * tmp3;
        mt = mw(end) + tmp2 * [ws_prev(1:end-1)-mw(1:end-1); 0] * tmp3;

        tmp1 = 1 - .5 * erfc(mt/realsqrt(2)/sqrt(st));
        if abs(tmp1) < 1e-9
            ws_prev(end) = 1e-121;
        else

```

```

        ws_prev(end) = mt + (tmp1)^-1 * (2*pi)^-(1/2) * sqrt(st) * exp(-
.5*mt^2/st);
        end
    end
    ws(:,q) = ws_prev;

    % compute the mean of gamma distribution of vita
    vita(q) = (M/2 + aVita + N/2) / (1/2 * norm(y-Phi*ws(:,q),2)^2 + bVita +
1/2 * ws(:,q).' * L * ws(:,q));

    % compute the mean of distribution of gamma
    gm(:,q) = sqrt(2*lamda(:,q-1)./pi).*exp(sqrt(lamda(:,q-
1).*vita(q)).*(ws(:,q))).*(vita(q)*ws(:,q).^2./lamda(:,q-
1)).^(3/4).*besselk(3/2,sqrt(lamda(:,q-1).*vita(q)).*ws(:,q)));
    L = diag(1./gm(:,q));

    % compute the mean of gamma distribution of lamda
    lamda(:,q) = (1 + aLamda)./(1/2 * gm(:,q) + bLamda);

    % Termination if necessary
    if ( mean( ( norm( ws(:,q) - ws(:,q-1), 2 ) .^ 2 ) ) < 10e-6 )
        break;
    end
end

w_BiICE = ws(:,q);
vita_fin = vita(q,1);
gm_fin = gm(:,q);
lamda_fin = lamda(:,q);

```

### **BiICE single lamda pixel correlation.m:**

```

function [ w_BiICE, vita_fin, gm_fin, lamda_fin, q ] =
BiICE_single_lamda_pixel_correlation ( Phi, y, MaxIter, w_pr, vita_pr, gm_pr,
lamda_pr, aVita, bVita, aLamda, bLamda )
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This code implements the fast Laplace algorithm from the following paper:
%
% [1] K. E. Themelis, A. A. Rontogiannis, K. D. Koutroumbas. "A Novel
Hierarchical Bayesian Approach for Sparse Semi-Supervised Hyperspectral
Unmixing," submitted for publication, IEEE Transactions on Signal Processing,
December 2010.
%
% Function:
%
% [ w_BiICE, vita_fin, gm_fin, lamda_fin, q ] =
BiICE_single_lamda_pixel_correlation ( Phi, y, MaxIter, w_pr, vita_pr, gm_pr,
lamda_pr, aVita, bVita, aLamda, bLamda )
%
% Inputs:
% PHI:          measurement matrix
% y:           vector of measurements
% MaxIter:     number of maximum iterations
% w_pr:       the final estimation of the correlated pixel
% vita_pr:    the final vita of the correlated pixel
% gm_pr:      the final gm of the correlated pixel
% lamda_pr:   the final lamda of the correlated pixel
% aVita:      the shape parameter of the gamma distribution of vita
% bVita:      the inverse of the scale parameter of the gamma distribution
%             of vita
% aLamda:     the shape parameter of the gamma distribution of lamda
% bLamda:     the inverse of the scale parameter of the gamma distribution

```

```

%                               of lamda
%
% Outputs:
%   w_BiICE:   final estimation vector
%   vita_fin:  conditional means for vita in the last step
%   gm_fin:    conditional means for gamma in the last step
%   lamda_fin: conditional means for lamda in the last step
%   q:         total number of iterations that were necessary for
%              convergence
%
% (c) 2010 K. E. Themelis, A. A. Rontogiannis, K. D. Koutroumbas
% Modified: (c) June 2012 Angelopoulos Vasileios, Kafouris Pavlos
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

[M,N] = size(Phi);

% Compute the covariance matrix
Phi_cov = Phi.' * Phi;

% Initialization
ws = ones(N,MaxIter); % Conditional means for w
ws(:,1) = w_pr;
gm = ones(N,MaxIter); % Conditional means for gamma
gm(:,1) = gm_pr;
L = diag(1./gm(:,1));
lamda = ones(1,MaxIter); % Conditional means for lambda
lamda(1,1) = lamda_pr;
vita = ones(MaxIter,1); % Conditional means for vita
vita(1,1) = vita_pr;

for q = 2 : MaxIter

    % Sample the gaussian distribution of w
    Sw = vita(q-1)^-1 * inv( Phi_cov + L);
    Sw_inv = vita(q-1) * ( Phi_cov + L);
    mw = vita(q-1) * Sw * Phi.' * y;
    ws_prev = mw;

    for t = 1 : N

        idx1 = cell(1);
        idx1{1} = mod((0:N-1)+1, N)+1;
        ws_prev = ws_prev(idx1{:});
        mw = mw(idx1{:});

        idx2 = {idx1{1} idx1{1}};
        Sw = Sw(idx2{:});
        Sw_inv = Sw_inv(idx2{:});
        S1 = Sw(1:end-1,end);

        tmp2 = [S1.' 0] * Sw_inv ;
        qalpha = tmp2 * [S1; 0];
        tmp3 = Sw(end,end) / (qalpha + Sw(end,end));
        st = Sw(end,end) - qalpha * tmp3;
        mt = mw(end) + tmp2 * [ws_prev(1:end-1)-mw(1:end-1); 0] * tmp3;

        tmp1 = 1 - .5 * erfc(mt/realsqrt(2)/sqrt(st));
        if abs(tmp1) < 1e-9
            ws_prev(end) = 1e-121;
        else
            ws_prev(end) = mt + (tmp1)^-1 * (2*pi)^-(1/2) * sqrt(st) * exp(-
.5*mt^2/st);
        end
    end
end

```

```

end
ws(:,q) = ws_prev;

% Compute the mean of gamma distribution of vita
vita(q) = (M/2 + aVita + N/2) / (1/2 * norm(y-Phi*ws(:,q),2)^2 + bVita +
1/2 * ws(:,q) .' * L * ws(:,q));

% Compute the mean of distribution of gamma
gm(:,q) = sqrt(2*lamda(1,q-1)./pi).*exp(sqrt(lamda(1,q-1).*vita(q)) .*
(ws(:,q))).*(vita(q)*ws(:,q).^2./lamda(1,q-1)) .^(3/4) .* bessellk (3/2, sqrt
(lamda(1,q-1).*vita(q)).*ws(:,q)));
L = diag(1./gm(:,q));

% Compute lamda
lamda(1,q) = (N + aLamda)./(sum(1/2 * gm(:,q)) + bLamda);

% Termination if necessary
if ( mean( ( norm( ws(:,q) - ws(:,q-1), 2 ) .^ 2 ) ) < 10e-6 )
    break;
end

end

w_BiICE = ws(:,q);
vita_fin = vita(q,1);
gm_fin = gm(:,q);
lamda_fin = lamda(1,q);

```

## Άλλες συναρτήσεις:

### matrix mutual coherence computation.m:

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% (c) June 2012 Angelopoulos Vasileios, Kafouris Pavlos %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

X = X / ( diag( sqrt( diag( X' * X ) ) ) );
mcoh = abs( X' * X );
V = diag( mcoh );
V = -diag( V, 0 );
mcoh = mcoh + V;
max( max( mcoh ) )

```

### randfixedsum.m:

```

% The current function is available in:
% http://www.mathworks.com/matlabcentral/fileexchange/9700/
% and was created by Roger Stafford in 19 Jan 2006

function [x,v] = randfixedsum(n,m,s,a,b)

% [x,v] = randfixedsum(n,m,s,a,b)
%
% This generates an n by m array x, each of whose m columns contains n
% random values lying in the interval [a,b], but subject to the condition that
% their sum be equal to s. The scalar value s must accordingly satisfy n*a <=
% s <= n*b. The distribution of values is uniform in the sense that it has
% the conditional probability distribution of a uniform distribution over the
% whole n-cube, given that the sum of the x's is s.

```

```

% The scalar v, if requested, returns with the total n-1 dimensional volume
%(content) of the subset satisfying this condition. Consequently if v,
%considered as a function of s and divided by sqrt(n), is integrated with
%respect to s from s = a to s = b, the result would necessarily be the n-
%dimensional volume of the whole cube, namely (b-a)^n.
% This algorithm does no "rejecting" on the sets of x's it obtains. It is
%designed to generate only those that satisfy all the above conditions and to
%do so with a uniform distribution.
% It accomplishes this by decomposing the space of all possible x sets
%(columns) into n-1 dimensional simplexes. (Line segments, triangles, and
%tetrahedra, are one-, two-, and three-dimensional examples of simplexes,
%respectively.) It makes use of three different sets of 'rand' variables,
%one to locate values uniformly within each type of simplex, another to
%randomly select representatives of each different type of simplex in
%proportion to their volume, and a third to perform random permutations to
%provide an even distribution of simplex choices among like types. For
%example, with n equal to 3 and s set at, say, 40% of the way from a towards
%b, there will be 2 different types of simplex, in this case triangles, each
%with its own area, and 6 different versions of each from permutations, for a
%total of 12 triangles, and these all fit together to form a particular
%planar non-regular hexagon in 3 dimensions, with v returned set equal to the
%hexagon's area.
%
% Roger Stafford - Jan. 19, 2006

% Check the arguments.
if (m~=round(m))|(n~=round(n))|(m<0)|(n<1)
    error('n must be a whole number and m a non-negative integer.')
elseif (s<n*a)|(s>n*b)|(a>=b)
    error('Inequalities n*a <= s <= n*b and a < b must hold.')
end

% Rescale to a unit cube: 0 <= x(i) <= 1
s = (s-n*a)/(b-a);

% Construct the transition probability table, t.
% t(i,j) will be utilized only in the region where j <= i + 1.
k = max(min(floor(s),n-1),0); % Must have 0 <= k <= n-1
s = max(min(s,k+1),k); % Must have k <= s <= k+1
s1 = s - [k:-1:k-n+1]; % s1 & s2 will never be negative
s2 = [k+n:-1:k+1] - s;
w = zeros(n,n+1); w(1,2) = realmax; % Scale for full 'double' range
t = zeros(n-1,n);
tiny = 2^(-1074); % The smallest positive matlab 'double' no.
for i = 2:n
    tmp1 = w(i-1,2:i+1).*s1(1:i)/i;
    tmp2 = w(i-1,1:i).*s2(n-i+1:n)/i;
    w(i,2:i+1) = tmp1 + tmp2;
    tmp3 = w(i,2:i+1) + tiny; % In case tmp1 & tmp2 are both 0,
    tmp4 = (s2(n-i+1:n) > s1(1:i)); % then t is 0 on left & 1 on right
    t(i-1,1:i) = (tmp2./tmp3).*tmp4 + (1-tmp1./tmp3).*(~tmp4);
end

% Derive the polytope volume v from the appropriate element in the bottom row
of w.
v = n^(3/2)*(w(n,k+2)/realmax)*(b-a)^(n-1);

% Now compute the matrix x.
x = zeros(n,m);
if m == 0, return, end % If m is zero, quit with x = []
rt = rand(n-1,m); % For random selection of simplex type
rs = rand(n-1,m); % For random location within a simplex
s = repmat(s,1,m);

```

```

j = repmat(k+1,1,m); % For indexing in the t table
sm = zeros(1,m); pr = ones(1,m); % Start with sum zero & product 1
for i = n-1:-1:1 % Work backwards in the t table
    e = (rt(n-i,:) <= t(i,j)); % Use rt to choose a transition
    sx = rs(n-i,:).^(1/i); % Use rs to compute next simplex coord.
    sm = sm + (1-sx).*pr.*s/(i+1); % Update sum
    pr = sx.*pr; % Update product
    x(n-i,:) = sm + pr.*e; % Calculate x using simplex coords.
    s = s - e; j = j - e; % Transition adjustment
end
x(n,:) = sm + pr.*s; % Compute the last x

% Randomly permute the order in the columns of x and rescale.
rp = rand(n,m); % Use rp to carry out a matrix 'randperm'
[ig,p] = sort(rp); % The values placed in ig are ignored
x = (b-a)*x(p+repmat([0:n*(m-1)],n,1))+a; % Permute & rescale x

return

```

## Συναρτήσεις και scripts πειραματικής μελέτης:

### single lamda experiment Phi.m:

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% (c) June 2012 Angelopoulos Vasileios, Kafouris Pavlos %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

close all
clear all
clc
load Phi

% Initializations
Iter = 2;
j = 1;
MaxIter = 100;
aVita = 0.01;
bVita = 0.01;
aLamda = 0.01;
bLamda = 0.01;
Num_of_patterns2 = 220;
MSE_lamda_vector = ones( 1, Iter );
MSE_lamda_vector_sparse = ones( 1, Iter );
MSE_single_lamda = ones( 1, Iter );
MSE_single_lamda_sparse = ones( 1, Iter );
gen_MSE_lamda_vector = ones( 1, 90 );
gen_MSE_lamda_vector_sparse = ones( 1, 90 );
gen_MSE_single_lamda = ones( 1, 90 );
gen_MSE_single_lamda_sparse = ones( 1, 90 );
vector_time = ones( 1, Iter );
single_time = ones( 1, Iter );

% Executing experiment using the signature matrix 'Phi'
disp( sprintf( 'Executing experiment using the signature matrix
"Phi"\n\n\n\n\n' ) );

for SNR_db = 10 : 5 : 50
    disp( sprintf( 'Executing experiment for SNR: %1.1f dB\n\n\n\n', SNR_db )
);

    SNR = 10 ^ ( SNR_db / 10 );

```

```

for Sparsity_level = 3 : 5 : 48
    disp( sprintf( 'Executing experiment for sparsity level: %d\n\n\n\n',
Sparsity_level ) );

    for i = 1 : Iter
        % Computation of the original abundance vector
        w_real = sprandn( Num_of_patterns2, 1, Sparsity_level./
Num_of_patterns2 );
        sp_rows = find( w_real );
        while( size( sp_rows ) ~= Sparsity_level )
            w_real = sprandn( Num_of_patterns2, 1, Sparsity_level./
Num_of_patterns2 );
            sp_rows = find( w_real );
        end
        w_real( sp_rows( : ) ) = randfixedsum( Sparsity_level, 1, 1, 0, 1
);

        % Computation of the AWG noise
        var = ( sum( ( Phi * w_real ) .^ 2 ) / length( Phi * w_real ) ) /
SNR;
        n = sqrt( var ) .* randn( 453, 1 );

        % Computation of the hyperspectral image pixel vector y
        y = Phi * w_real + n;

        tic;
        % Execution of the Bi-ICE Algorithm with lamda vector on the
pixel vector
        w_BiICE_vector = BiICE ( Phi, y, MaxIter, aVita, bVita, aLamda,
bLamda );
        % Compute the elapsed time for the previous execution
        vector_time( i ) = toc;

        % Computation of the Mean Square Error after the execution of the
original algorithm for the whole w, but also for the sparse part of it
        MSE_lamda_vector( i ) = mean( ( norm( ( w_real - w_BiICE_vector
), 2 ) .^ 2 ) ./ ( norm( w_real, 2 ) .^ 2 ) );
        MSE_lamda_vector_sparse( i ) = mean( ( norm( ( w_real( sp_rows( :
) ) - w_BiICE_vector( sp_rows( : ) ) ), 2 ) .^ 2 ) ./ ( norm( w_real(
sp_rows( : ) ), 2 ) .^ 2 ) );

        tic
        % Execution of the Bi-ICE Algorithm with single lamda on the
pixel vector
        w_BiICE_single = BiICE_single_lamda ( Phi, y, MaxIter, aVita,
bVita, aLamda, bLamda );
        % Compute the elapsed time for the previous execution
        single_time( i ) = toc;

        % Computation of the Mean Square Error after the execution of the
variant algorithm for the whole w, but also for the sparse part of it
        MSE_single_lamda( i ) = mean( ( norm( ( w_real - w_BiICE_single
), 2 ) .^ 2 ) ./ ( norm( w_real, 2 ) .^ 2 ) );
        MSE_single_lamda_sparse( i ) = mean( ( norm( ( w_real( sp_rows( :
) ) - w_BiICE_single( sp_rows( : ) ) ), 2 ) .^ 2 ) ./ ( norm( w_real(
sp_rows( : ) ), 2 ) .^ 2 ) );

    end

    % Printing on screen the results of this subpart of the experiment
    disp( sprintf( 'Mean MSE for lamda vector for whole w: %f', mean(
MSE_lamda_vector ) ) );

```

```

        disp( sprintf( 'Mean MSE for lamda vector for sparse part of w: %f',
mean( MSE_lamda_vector_sparse ) ) );
        disp( sprintf( 'Mean MSE for single lamda for whole w: %f', mean(
MSE_single_lamda ) ) );
        disp( sprintf( 'Mean MSE for single lamda for sparse part of w:
%f\n', mean( MSE_single_lamda_sparse ) ) );

        % Printing on screen the mean required time for each subpart of the
experiment
        disp( sprintf( 'Total elapsed time for executing the algorithm with
lamda vector: %f seconds', sum( vector_time ) ) );
        disp( sprintf( 'Mean elapsed time for executing the algorithm with
lamda vector: %f seconds', mean( vector_time ) ) );
        disp( sprintf( 'Total elapsed time for executing the algorithm with
single lamda: %f seconds', sum( single_time ) ) );
        disp( sprintf( 'Mean elapsed time for executing the algorithm with
single lamda: %f seconds\n\n', mean( single_time ) ) );

        % Saving results of this step for generalized statistics
gen_MSE_lamda_vector( j ) = mean( MSE_lamda_vector );
gen_MSE_lamda_vector_sparse( j ) = mean( MSE_lamda_vector_sparse );
gen_MSE_single_lamda( j ) = mean( MSE_single_lamda );
gen_MSE_single_lamda_sparse( j ) = mean( MSE_single_lamda_sparse );

        j = j + 1;
    end
    disp( sprintf( '\n\n\n\n' ) );
end

disp( sprintf( '\n\n\n\n\n\n' ) );

% Printing on screen the results of this subpart of the experiment
disp( sprintf( 'Total mean MSE for lamda vector for whole w after whole
experiment: %f', mean( gen_MSE_lamda_vector ) ) );
disp( sprintf( 'Total mean MSE for lamda vector for sparse part of w after
whole experiment: %f', mean( gen_MSE_lamda_vector_sparse ) ) );
disp( sprintf( 'Total mean MSE for single lamda for whole w after whole
experiment: %f', mean( gen_MSE_single_lamda ) ) );
disp( sprintf( 'Total mean MSE for single lamda for sparse part of w after
whole experiment: %f\n', mean( gen_MSE_single_lamda_sparse ) ) );

```

## **w experiment Phi.m:**

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% (c) June 2012 Angelopoulos Vasileios, Kafouris Pavlos %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

close all
clear all
clc
load Phi;

Num_of_patterns2 = 220;
MaxIter = 100;
aVita = 0.01;
bVita = 0.01;
aLamda = 0.01;
bLamda = 0.01;
w_real_for_plot = ones( Num_of_patterns2, MaxIter );
Sparsity_level = 3;
SNR_db = 25;

```



```

% Creation of the original abundance vector
w_real = zeros( Num_of_patterns2, 1 );
sp_rows = [ 28, 72, 186 ];
w_real( sp_rows ) = [ 0.1397, 0.2305, 0.6298 ];

% Computation of the hyperspectral image pixel vector y
SNR = 10 ^ ( SNR_db / 10 );
var = ( sum( ( Phi * w_real ) .^ 2 ) / length( Phi * w_real ) ) / SNR;
n = sqrt( var ) .* randn( 453, 1 );

y = Phi * w_real + n;

% Execution of the Bi-ICE Algorithm with lamda vector on the pixel vector
[ w_BiICE_vector, ws_vector, ~, ~, ~ ] = BiICE ( Phi, y, MaxIter, aVita,
bVita, aLamda, bLamda );

% Execution of the Bi-ICE Algorithm with single lamda on the pixel vector
[ w_BiICE_single, ws_single, ~, ~, ~ ] = BiICE_single_lamda ( Phi, y,
MaxIter, aVita, bVita, aLamda, bLamda );

% Computation of the Mean Square Error after the execution of the original
algorithm for the whole w, but also for the sparse part of it
MSE_lamda_vector = mean( ( norm( ( w_real - w_BiICE_vector ), 2 ) .^ 2 ) ./ (
norm( w_real, 2 ) .^ 2 ) );
MSE_lamda_vector_sparse = mean( ( norm( ( w_real( sp_rows( : ) ) -
w_BiICE_vector( sp_rows( : ) ) ), 2 ) .^ 2 ) ./ ( norm( w_real( sp_rows( : )
), 2 ) .^ 2 ) );

% Computation of the Mean Square Error after the execution of the variant
algorithm for the whole w, but also for the sparse part of it
MSE_single_lamda = mean( ( norm( ( w_real - w_BiICE_single ), 2 ) .^ 2 ) ./ (
norm( w_real, 2 ) .^ 2 ) );
MSE_single_lamda_sparse = mean( ( norm( ( w_real( sp_rows( : ) ) -
w_BiICE_single( sp_rows( : ) ) ), 2 ) .^ 2 ) ./ ( norm( w_real( sp_rows( : )
), 2 ) .^ 2 ) );

% Printing on screen the results of the experiment
disp( sprintf( '\nMSE for lamda vector for whole w: %f', MSE_lamda_vector ) );
disp( sprintf( 'MSE for lamda vector for sparse part of w: %f',
MSE_lamda_vector_sparse ) );
disp( sprintf( 'MSE for single lamda for whole w: %f', MSE_single_lamda ) );
disp( sprintf( 'MSE for single lamda for sparse part of w: %f\n',
MSE_single_lamda_sparse ) );

for i = 1 : Num_of_patterns2
    w_real_for_plot(i,:) = w_real(i) .* w_real_for_plot(i,:);
end

figure(1);
axis([0 100 0 1]);
title( 'Φ - BI-ICE' );
xlabel('Number of Iterations');
ylabel('Abundances');

for i = 1 : Num_of_patterns2
    hold on;
    plot( ws_vector(i,:) );
end
for i = 1 : Sparsity_level
    hold on;
    plot( w_real_for_plot(sp_rows(i),:), '--', 'Color', 'r' );
end

```

```

figure(2);
axis([0 100 0 1]);
title( 'Φ - BI-ICE-single' );
xlabel('Number of Iterations');
ylabel('Abundances');

for i = 1 : Num_of_patterns2
    hold on;
    plot( ws_single(i,:) );
end
for i = 1 : Sparsity_level
    hold on;
    plot( w_real_for_plot(sp_raws(i,:),), '--', 'Color', 'r' );
end

```

### single lamda experiment plots SNR 20.m:

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% (c) June 2012 Angelopoulos Vasileios, Kafouris Pavlos %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

figure(1);
semilogy( gen_MSE_lamda_vector( 21 : 30 ), '-o', 'MarkerFaceColor', 'blue' )
hold all;
semilogy( gen_MSE_single_lamda( 21 : 30 ), '--rs', 'Color', 'r',
'MarkerFaceColor', 'r' )
set( gca, 'XTick', 1 : 10 )
set( gca, 'XTickLabel', {'3','8','13','18','23','28','33','38','43','48'} )
title( 'Φ - all abundances' );
xlabel('ξ');
ylabel('MSE');
hleg1 = legend( 'BI-ICE', 'BI-ICE-single' );
set( hleg1, 'Location', 'NorthWest' );

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
figure(2);
semilogy( gen_MSE_lamda_vector_sparse( 21 : 30 ), '-o', 'MarkerFaceColor',
'blue' )
hold all;
semilogy( gen_MSE_single_lamda_sparse( 21 : 30 ), '--rs', 'Color', 'r',
'MarkerFaceColor', 'r' )
set( gca, 'XTick', 1 : 10 )
set( gca, 'XTickLabel', {'3','8','13','18','23','28','33','38','43','48'} )
title( 'Φ - sparse abundances' );
xlabel('ξ');
ylabel('MSE');
hleg1 = legend( 'BI-ICE', 'BI-ICE-single' );
set( hleg1, 'Location', 'NorthWest' );

```

### single lamda experiment plots ksi 8.m:

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% (c) June 2012 Angelopoulos Vasileios, Kafouris Pavlos %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

figure(1);
semilogy( gen_MSE_lamda_vector( 2 : 10 : 90 ), '-o', 'MarkerFaceColor',
'blue' )
hold all;

```

```
semilogy( gen_MSE_single_lamda( 2 : 10 : 90 ), '--rs', 'Color', 'r',
'MarkerFaceColor', 'r' )
set( gca, 'XTick', 1 : 9 )
set( gca, 'XTickLabel', {'10','15','20','25','30','35','40','45','50'} )
title( 'Φ - all abundances' );
xlabel('SNR');
ylabel('MSE');
hleg1 = legend( 'BI-ICE', 'BI-ICE-single' );
set( hleg1, 'Location', 'NorthWest' );
```

%%%

```
figure(2);
semilogy( gen_MSE_lamda_vector_sparse( 2 : 10 : 90 ), '-o',
'MarkerFaceColor', 'blue' )
hold all;
semilogy( gen_MSE_single_lamda_sparse( 2 : 10 : 90 ), '--rs', 'Color', 'r',
'MarkerFaceColor', 'r' )
set( gca, 'XTick', 1 : 9 )
set( gca, 'XTickLabel', {'10','15','20','25','30','35','40','45','50'} )
title( 'Φ - sparse abundances' );
xlabel('SNR');
ylabel('MSE');
hleg1 = legend( 'BI-ICE', 'BI-ICE-single' );
set( hleg1, 'Location', 'NorthWest' );
```

### MSE computation.m:

%%%  
 % (c) June 2012 Angelopoulos Vasileios, Kafouris Pavlos %  
 %%

```
MSE_vec = ones( 250, 191 );
MSE_sin = ones( 250, 191 );

for i = 1 : 250
    for j = 1 : 191
        MSE_vec( i, j ) = mean( ( norm( ( W_vec( :, i, j ) - W_vec_cor( :,
i, j ), 2 ) .^ 2 ) ./ ( norm( W_vec( :, i, j ), 2 ) .^ 2 ) );
        MSE_sin( i, j ) = mean( ( norm( ( W_sin( :, i, j ) - W_sin_cor( :,
i, j ), 2 ) .^ 2 ) ./ ( norm( W_sin( :, i, j ), 2 ) .^ 2 ) );
    end
end

MSE_vec_total = mean( mean( MSE_vec ) )
MSE_sin_total = mean( mean( MSE_sin ) )
```

### num of iterations plots.m:

%%%  
 % (c) June 2012 Angelopoulos Vasileios, Kafouris Pavlos %  
 %%

```
figure(1);
semilogy( Q_sin_cor, '-' )
hold all
semilogy( Q_sin, '--', 'Color', 'r' )
title( 'BI-ICE-single' );
xlabel('Processed pixels');
```

```

ylabel('Total number of iterations');
hleg1 = legend( 'BI-ICE-single with spatial correlation', 'BI-ICE-single
without spatial correlation' );
set( hleg1, 'Location', 'NorthWest' );

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

figure(2);
semilogy( Q_vec_cor, '-' )
hold all
semilogy( Q_vec, '--', 'Color', 'r' )
title( 'BI-ICE' );
xlabel('Processed pixels');
ylabel('Total number of iterations');
hleg2 = legend( 'BI-ICE with spatial correlation', 'BI-ICE without spatial
correlation' );
set( hleg2, 'Location', 'NorthWest' );

```

**image with BiICE.m:**

```

function [ w_BiICE_im, Q ] = image_with_BiICE ( Im, Phi, MaxIter, aVita,
bVita, aLamda, bLamda )
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Function:
%
% [ w_BiICE_im, Q ] = image_with_BiICE ( Im, Phi, MaxIter, aVita, bVita,
aLamda, bLamda )
%
% Inputs:
%   Im:           the matrix that represents the whole image
%   Phi:          measurement matrix
%   MaxIter:      number of maximum iterations
%   aVita:        the shape parameter of the gamma distribution of vita
%   bVita:        the inverse of the scale parameter of the gamma distribution
%                 of vita
%   aLamda:       the shape parameter of the gamma distribution of lamda
%   bLamda:       the inverse of the scale parameter of the gamma distribution
%                 of lamda
%
% Outputs:
%   w_BiICE_im:  a NxMxNum_of_patterns array that represents the
%                 final estimation vectors for every pixel of the image
%   Q:           a NxM+1 vector that represents the total number of
%                 iterations that the algorithm has done until the specific
number of pixel
%
% (c) June 2012 Angelopoulos Vasileios, Kafouris Pavlos
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

Num_of_patterns = 14;

[ ~, N, M ] = size( Im );

% Initialization
w_BiICE_im = ones( Num_of_patterns, N, M );
Q = zeros( N*M + 1, 1 );
k = 2;

tic

for i = 1 : N

```

```

for j = 1 : M

    % Execution of the Bi-ICE Algorithm with single lamda with pixel
    correlation on the pixel vector
    [ w_BiICE, ~, ~, ~, ~, q ] = BiICE ( Phi, Im( :, i, j ), MaxIter,
    aVita, bVita, aLamda, bLamda );

    % Preparation for next step
    w_BiICE_im( :, i, j ) = w_BiICE;
    Q( k ) = Q( k - 1 ) + q;
    k = k + 1;
end
end

total_time = toc;

% Printing on screen the required time for the whole experiment
disp( sprintf( '\nElapsed time for executing the algorithm with lamda vector
without pixel correlation: %f seconds', total_time ) );
disp( sprintf( 'Mean time for executing the algorithm for each pixel: %f
seconds\n', total_time ./ ( N * M ) ) );

w2 = squeeze( w_BiICE_im(2, :, :) );
w2 = w2 - min( min( w2 ) );
w2 = w2 ./ max( max( w2 ) );
w2 = round( w2 * 255 );
w2 = uint8( w2 );

w8 = squeeze( w_BiICE_im(8, :, :) );
w8 = w8 - min( min( w8 ) );
w8 = w8 ./ max( max( w8 ) );
w8 = round( w8 * 255 );
w8 = uint8( w8 );

w1 = squeeze( w_BiICE_im(1, :, :) );
w1 = w1 - min( min( w1 ) );
w1 = w1 ./ max( max( w1 ) );
w1 = round( w1 * 255 );
w1 = uint8( w1 );

w14 = squeeze( w_BiICE_im(14, :, :) );
w14 = w14 - min( min( w14 ) );
w14 = w14 ./ max( max( w14 ) );
w14 = round( w14 * 255 );
w14 = uint8( w14 );

figure(1);
subplot(1,2,1);imagesc(w2,[0 255]);axis('image');axis('off');
subplot(1,2,2);imagesc(w8,[0 255]);axis('image');axis('off');
figure(2);
subplot(1,2,1);imagesc(w1,[0 255]);axis('image');axis('off');
subplot(1,2,2);imagesc(w14,[0 255]);axis('image');axis('off');

```

### image with BiICE single.m:

```

function [ w_BiICE_im, Q ] = image_with_BiICE_single ( Im, Phi, MaxIter,
aVita, bVita, aLamda, bLamda )
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Function:
% [ w_BiICE_im, Q ] = image_with_BiICE_single ( Im, Phi, MaxIter, aVita,
bVita, aLamda, bLamda )

```

```

% Inputs:
%   Im:           the matrix that represents the whole image
%   Phi:          measurement matrix
%   MaxIter:      number of maximum iterations
%   aVita:        the shape parameter of the gamma distribution of vita
%   bVita:        the inverse of the scale parameter of the gamma distribution
%                 of vita
%   aLamda:       the shape parameter of the gamma distribution of lamda
%   bLamda:       the inverse of the scale parameter of the gamma distribution
%                 of lamda
%
% Outputs:
%   w_BiICE_im:  a NxMxNum_of_patterns array that represents the
%                 final estimation vectors for every pixel of the image
%   Q:           a NxM+1 vector that represents the total number of
%                 iterations that the algorithm has done until the specific
number of pixel
%
% (c) June 2012 Angelopoulos Vasileios, Kafouris Pavlos
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

Num_of_patterns = 14;

[ ~, N, M ] = size( Im );

% Initialization
w_BiICE_im = ones( Num_of_patterns, N, M );
Q = zeros( N*M + 1, 1 );
k = 2;

tic

for i = 1 : N
    for j = 1 : M

        % Execution of the Bi-ICE Algorithm with single lamda with pixel
correlation on the pixel vector
        [ w_BiICE, ~, ~, ~, ~, q ] = BiICE_single_lamda ( Phi, Im( :, i, j ),
MaxIter, aVita, bVita, aLamda, bLamda );

        % Preparation for next step
        w_BiICE_im( :, i, j ) = w_BiICE;
        Q( k ) = Q( k - 1 ) + q;
        k = k + 1;
    end
end

total_time = toc;

% Printing on screen the required time for the whole experiment
disp( sprintf( '\nElapsed time for executing the algorithm with single lamda
without pixel correlation: %f seconds', total_time ) );
disp( sprintf( 'Mean time for executing the algorithm for each pixel: %f
seconds\n', total_time ./ ( N * M ) ) );

w2 = squeeze( w_BiICE_im(2, :, :) );
w2 = w2 - min( min( w2 ) );
w2 = w2 ./ max( max( w2 ) );
w2 = round( w2 * 255 );
w2 = uint8( w2 );

w8 = squeeze( w_BiICE_im(8, :, :) );
w8 = w8 - min( min( w8 ) );

```

```
w8 = w8 ./ max( max( w8 ) );
w8 = round( w8 * 255 );
w8 = uint8( w8 );

w1 = squeeze( w_BiICE_im(1, :, :) );
w1 = w1 - min( min( w1 ) );
w1 = w1 ./ max( max( w1 ) );
w1 = round( w1 * 255 );
w1 = uint8( w1 );

w14 = squeeze( w_BiICE_im(14, :, :) );
w14 = w14 - min( min( w14 ) );
w14 = w14 ./ max( max( w14 ) );
w14 = round( w14 * 255 );
w14 = uint8( w14 );

figure(1);
subplot(1,2,1);imagesc(w2,[0 255]);axis('image');axis('off');
subplot(1,2,2);imagesc(w8,[0 255]);axis('image');axis('off');
figure(2);
subplot(1,2,1);imagesc(w1,[0 255]);axis('image');axis('off');
subplot(1,2,2);imagesc(w14,[0 255]);axis('image');axis('off');
```

### **lamda vector pixel correlation experiment.m:**

```
function [ w_BiICE_im, Q ] = lamda_vector_pixel_correlation_experiment ( N,
M, Im, Phi, MaxIter, aVita, bVita, aLamda, bLamda )
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Function:
%
% [ w_BiICE_im, Q ] = lamda_vector_pixel_correlation_experiment ( N, M, Im,
Phi, MaxIter, aVita, bVita, aLamda, bLamda )
%
% Inputs:
% N:          the image length
% M:          the image width
% Im:         the matrix that represents the whole image
% Phi:        measurement matrix
% MaxIter:    number of maximum iterations
% aVita:      the shape parameter of the gamma distribution of vita
% bVita:      the inverse of the scale parameter of the gamma distribution
%             of vita
% aLamda:     the shape parameter of the gamma distribution of lamda
% bLamda:     the inverse of the scale parameter of the gamma distribution
%             of lamda
%
% Outputs:
% w_BiICE_im: a NxMxNum_of_patterns array that represents the
%             final estimation vectors for every pixel of the image
% Q:          a NxM+1 vector that represents the total number of
%             iterations that the algorithm has done until the specific
number of pixel
%
% (c) June 2012 Angelopoulos Vasileios, Kafouris Pavlos
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

Num_of_patterns = 14;

% Initialization
w_BiICE_im = ones( Num_of_patterns, N, M );
Q = zeros( N*M + 1, 1 );
```

```

w_pr = ones( Num_of_patterns, 1 );
gm_pr_line = ones( Num_of_patterns, 1 );
vita_pr_line = 1;
lamda_pr_line = ones( Num_of_patterns, 1 );
k = 2;

tic

for i = 1 : N

    % Preparation for next row
    if i ~= 1
        w_pr = w_BiICE_im( :, i - 1, 1 );
    end
    gm_pr = gm_pr_line;
    vita_pr = vita_pr_line;
    lamda_pr = lamda_pr_line;

    for j = 1 : M

        % Execution of the Bi-ICE Algorithm with single lamda with pixel
        correlation on the pixel vector
        [ w_BiICE, vita_fin, gm_fin, lamda_fin, q ] = BiICE_pixel_correlation
        ( Phi, Im( :, i, j ), MaxIter, w_pr, vita_pr, gm_pr, lamda_pr, aVita, bVita,
        aLamda, bLamda );

        % Preparation for next step
        w_BiICE_im( :, i, j ) = w_BiICE;
        Q( k ) = Q( k - 1 ) + q;
        w_pr = w_BiICE;
        gm_pr = gm_fin;
        vita_pr = vita_fin;
        lamda_pr = lamda_fin;
        k = k + 1;

        if j == 1
            gm_pr_line = gm_fin;
            vita_pr_line = vita_fin;
            lamda_pr_line = lamda_fin;
        end
    end
end

total_time = toc;

% Printing on screen the required time for the whole experiment
disp( sprintf( '\nElapsed time for executing the algorithm with lamda vector
and pixel correlation: %f seconds', total_time ) );
disp( sprintf( 'Mean time for executing the algorithm for each pixel: %f
seconds\n', total_time ./ ( N * M ) ) );

w2 = squeeze( w_BiICE_im(2, :, :) );
w2 = w2 - min( min( w2 ) );
w2 = w2 ./ max( max( w2 ) );
w2 = round( w2 * 255 );
w2 = uint8( w2 );

w8 = squeeze( w_BiICE_im(8, :, :) );
w8 = w8 - min( min( w8 ) );
w8 = w8 ./ max( max( w8 ) );
w8 = round( w8 * 255 );
w8 = uint8( w8 );

```



```

w1 = squeeze( w_BiICE_im(1, :, :) );
w1 = w1 - min( min( w1 ) );
w1 = w1 ./ max( max( w1 ) );
w1 = round( w1 * 255 );
w1 = uint8( w1 );

w14 = squeeze( w_BiICE_im(14, :, :) );
w14 = w14 - min( min( w14 ) );
w14 = w14 ./ max( max( w14 ) );
w14 = round( w14 * 255 );
w14 = uint8( w14 );

figure(1);
subplot(1,2,1);imagesc(w2,[0 255]);axis('image');axis('off');
subplot(1,2,2);imagesc(w8,[0 255]);axis('image');axis('off');
figure(2);
subplot(1,2,1);imagesc(w1,[0 255]);axis('image');axis('off');
subplot(1,2,2);imagesc(w14,[0 255]);axis('image');axis('off');

```

### single lamda pixel correlation experiment.m:

```

function [ w_BiICE_im, Q ] = single_lamda_pixel_correlation_experiment ( N,
M, Im, Phi, MaxIter, aVita, bVita, aLamda, bLamda )
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Function:
%
% [ w_BiICE_im, Q ] = single_lamda_pixel_correlation_experiment ( N, M, Im,
Phi, MaxIter, aVita, bVita, aLamda, bLamda )
%
% Inputs:
% N:          the image length
% M:          the image width
% Im:         the matrix that represents the whole image
% Phi:        measurement matrix
% MaxIter:    number of maximum iterations
% aVita:      the shape parameter of the gamma distribution of vita
% bVita:      the inverse of the scale parameter of the gamma distribution
%             of vita
% aLamda:     the shape parameter of the gamma distribution of lamda
% bLamda:     the inverse of the scale parameter of the gamma distribution
%             of lamda
%
% Outputs:
% w_BiICE_im: a NxMxNum_of_patterns array that represents the
%             final estimation vectors for every pixel of the image
% Q:          a NxM+1 vector that represents the total number of
%             iterations that the algorithm has done until the specific
%             number of pixel
%
% (c) June 2012 Angelopoulos Vasileios, Kafouris Pavlos
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

Num_of_patterns = 14;

% Initialization
w_BiICE_im = ones( Num_of_patterns, N, M );
Q = zeros( N*M + 1, 1 );
w_pr = ones( Num_of_patterns, 1 );
gm_pr_line = ones( Num_of_patterns, 1 );
vita_pr_line = 1;
lamda_pr_line = 1;
k = 2;
B. Αγγελόπουλος, Π. Καφούρης

```

```

tic

for i = 1 : N

    % Preparation for next row
    if i ~= 1
        w_pr = w_BiICE_im( :, i - 1, 1 );
    end
    gm_pr = gm_pr_line;
    vita_pr = vita_pr_line;
    lamda_pr = lamda_pr_line;

    for j = 1 : M

        % Execution of the Bi-ICE Algorithm with single lamda with pixel
        correlation on the pixel vector
        [ w_BiICE, vita_fin, gm_fin, lamda_fin, q ] =
        BiICE_single_lamda_pixel_correlation ( Phi, Im( :, i, j ), MaxIter, w_pr,
        vita_pr, gm_pr, lamda_pr, aVita, bVita, aLamda, bLamda );

        % Preparation for next step
        w_BiICE_im( :, i, j ) = w_BiICE;
        Q( k ) = Q( k - 1 ) + q;
        w_pr = w_BiICE;
        gm_pr = gm_fin;
        vita_pr = vita_fin;
        lamda_pr = lamda_fin;
        k = k + 1;

        if j == 1
            gm_pr_line = gm_fin;
            vita_pr_line = vita_fin;
            lamda_pr_line = lamda_fin;
        end
    end
end

total_time = toc;

% Printing on screen the required time for the whole experiment
disp( sprintf( '\nElapsed time for executing the algorithm with single lamda
and pixel correlation: %f seconds', total_time ) );
disp( sprintf( 'Mean time for executing the algorithm for each pixel: %f
seconds\n', total_time ./ ( N * M ) ) );

w2 = squeeze( w_BiICE_im(2, :, :) );
w2 = w2 - min( min( w2 ) );
w2 = w2 ./ max( max( w2 ) );
w2 = round( w2 * 255 );
w2 = uint8( w2 );

w8 = squeeze( w_BiICE_im(8, :, :) );
w8 = w8 - min( min( w8 ) );
w8 = w8 ./ max( max( w8 ) );
w8 = round( w8 * 255 );
w8 = uint8( w8 );

w1 = squeeze( w_BiICE_im(1, :, :) );
w1 = w1 - min( min( w1 ) );
w1 = w1 ./ max( max( w1 ) );
w1 = round( w1 * 255 );
w1 = uint8( w1 );

```

Ανάπτυξη και αξιολόγηση αλγορίθμων συμπερασμού κατά Bayes κατάλληλων για φασματικό διαχωρισμό

```
w14 = squeeze( w_BiICE_im(14, :, :) );  
w14 = w14 - min( min( w14 ) );  
w14 = w14 ./ max( max( w14 ) );  
w14 = round( w14 * 255 );  
w14 = uint8( w14 );  
  
figure(1);  
subplot(1,2,1);imagesc(w2,[0 255]);axis('image');axis('off');  
subplot(1,2,2);imagesc(w8,[0 255]);axis('image');axis('off');  
figure(2);  
subplot(1,2,1);imagesc(w1,[0 255]);axis('image');axis('off');  
subplot(1,2,2);imagesc(w14,[0 255]);axis('image');axis('off');
```

## ΑΝΑΦΟΡΕΣ

- [1] K.E. Themelis, A.A. Rontogiannis, K.D. Koutroumbas, "A Novel Hierarchical Bayesian Approach for Sparse Semi-Supervised Hyperspectral Unmixing", *IEEE Transactions on Signal Processing*, vol. 60, n0. 2, pp. 585-599, Feb. 2012.
- [2] D. Landgrebe, "Hyperspectral image data analysis," *IEEE Signal Process. Mag.*, vol. 19, pp. 17–28, Jan. 2002.
- [3] G. Shaw and D. Manolakis, "Signal processing for hyperspectral image exploitation," *IEEE Signal Process. Mag.*, vol. 19, pp. 12–16, Jan. 2002.
- [4] N.Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Trans. Signal Process.*, vol. 19, pp. 44–57, Jan. 2002.
- [5] K. Themelis, A.A. Rontogiannis, "A Soft Constrained MAP Estimator for Supervised Hyperspectral Signal Unmixing", In *Proceedings of the 16th European Signal Processing Conference (EUSIPCO)*, Lausanne, Aug. 2008.
- [6] D. Manolakis, C. Siracusa, and G. Shaw, "Hyperspectral Subpixel Target Detection Using the Linear Mixing Model," *IEEE Trans. Geosci. Remote Sensing*, vol. 39, pp. 1392–1409, July 2001.
- [7] D. C. Heinz and C. I. Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, pp. 529–545, Mar. 2001.
- [8] H. Kwon and N. M. Nasrabadi, "Kernel orthogonal subspace projection for hyperspectral signal classification," *IEEE Trans. Geosci. Remote Sensing*, vol. 43, pp. 2952–2962, Dec. 2005.
- [9] C. I. Chang, "Orthogonal subspace projection (OSP) revisited: a comprehensive study and analysis," *IEEE Trans. Geosci. Remote Sensing*, vol. 43, pp. 502–518, March 2005.
- [10] J. M. Nascimento and J. M. Bioucas-Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, pp. 898–910, Apr. 2005.
- [11] S. Jia and Y. Qian, "Spectral and Spatial Complexity-Based Hyperspectral Unmixing," *IEEE Trans. Geosci. Remote Sensing*, vol. 45, pp. 3867–3879, Dec. 2007.
- [12] M. E. Winter, "N-FINDR: An algorithm for fast autonomous spectral end-member determination in hyperspectral data," *Proc. SPIE Imaging Spectrometry V*, vol. 3753, pp. 266–275, Jul. 1999.
- [13] N. Dobigeon, J.-Y. Tournier, and C.-I. Chang, "Semisupervised linear spectral unmixing using a hierarchical Bayesian model for hyperspectral imagery," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 2684–2695, Jul. 2008.
- [14] K. Themelis, A. A. Rontogiannis, and K. D. Koutroumbas, "Semisupervised hyperspectral unmixing via the weighted Lasso," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP'10)*, Dallas, TX, Mar. 2010.
- [15] J. W. Boardman, "Automating spectral unmixing of AVIRIS data using convex geometry concepts," in *Proc. Summaries 4th Ann. JPL Airborne Geosci. Workshop*, Wash., DC, 1993, vol. 1, pp. 11–14.
- [16] T. F. Coleman and Y. Li, "A reflective newton method for minimizing a quadratic function subject to bounds on some of the variables," *SIAM J. Optimiz.*, vol. 6, pp. 1040–1058, 1996.
- [17] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least angle regression," *Ann. Statist.*, vol. 32, pp. 407–499, Feb. 2002.

- [18] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, pp. 4655–4666, Dec. 2007.
- [19] D. L. Donoho, Y. Tsaig, I. Drori, and J. L. Starck, *Sparse Solution of Underdetermined Linear Equations by Stagewise Orthogonal Matching Pursuit* Dep. Statist., Stanford Univ., CA, 2006.
- [20] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2346–2356, Jun. 2008.
- [21] M. Figueiredo, "Adaptive sparseness for supervised learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 9, pp. 1150–1159, Sep. 2003.
- [22] N. Dobigeon, A. Hero, and J.-Y. Tournet, "Hierarchical Bayesian sparse image reconstruction with application to MRFM," *IEEE Trans. Image Process.*, vol. 18, no. 9, pp. 2059–2070, Sep. 2009.
- [23] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 721–741, Nov. 1984.
- [24] S. Babacan, R. Molina, and A. Katsaggelos, "Bayesian compressive sensing using Laplace priors," *IEEE Trans. Image Process.*, vol. 19, no. 1, pp. 53–63, Jan. 2010.
- [25] T. Park and C. George, "The Bayesian Lasso," *J. Amer. Statist. Assoc.*, vol. 103, no. 482, pp. 681–686, Jun. 2008.
- [26] H. Zou, "The adaptive Lasso and its oracle properties," *J. Amer. Statist. Assoc.*, vol. 101, pp. 1418–1429, Dec. 2006.
- [27] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, pp. 211–244, 2001.
- [28] R. Tibshirani, "Regression shrinkage and selection via the Lasso," *J. Royal Statist. Soc.*, vol. 58, no. 1, pp. 267–288, 1996.
- [29] D. F. Andrews and C. L. Mallows, "Scale mixtures of normal distributions," *J. Royal Statist. Soc., Ser. B*, vol. 36, no. 1, pp. 99–102, 1974.
- [30] J. Bioucas-Dias, "Bayesian wavelet-based image deconvolution: a GEM algorithm exploiting a class of heavy-tailed priors," *IEEE Trans. Image Process.*, vol. 15, no. 4, pp. 937–951, Apr. 2006.
- [31] M. Kyung, J. Gilly, M. Ghoshz, and G. Casella, "Penalized regression, standard errors, and Bayesian Lassos," *Bayesian Anal.*, vol. 5, pp. 369–412, Feb. 2010.
- [32] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [33] I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series, and Products*. New York: Academic, 1980.
- [34] N. L. Johnson and S. Kotz, *Continuous Univariate Distributions-1*. New York: Wiley, 1970.
- [35] C. P. Robert, "Simulation of truncated normal variables," *Statist. Comput.*, vol. 5, pp. 121–125, 1995.
- [36] G. Rodriguez-Yam, R. Davis, and L. Scharf, "Efficient Gibbs sampling of truncated multivariate normal with application to constrained linear regression," *Columbia Univ., New York*, 2004.
- [37] J. Bioucas-Dias and M. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *Proc. IEEE Int. Workshop on Hyperspectral Image and Signal Process.: Evolution in Remote Sens. (WHISPERS'10)*, Reykjavik, Iceland, Jun. 2010.
- [38] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2014–2039, Jun. 2011.

- [39] A. Bruckstein, M. Elad, and M. Zibulevsky, "On the uniqueness of nonnegative sparse solutions to underdetermined systems of equations," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 4813–4820, Nov. 2008.
- [40] G. H. Golub and C. F. Van Loan, *Matrix Computations* (Johns Hopkins Studies in Mathematical Sciences), 3rd ed. Baltimore, MD: The Johns Hopkins Univ. Press, 1996.
- [41] R. N. Clark, G. A. Swayze, R. Wise, K. E. Livo, T. M. Hoefen, R. F. Kokaly, and S. J. Sutley, *USGS Digital Spectral Library, 2007* [Online]. Available: <http://speclab.cr.usgs.gov/spectral.lib06/ds231/datatable.html>
- [42] *AVIRIS Free Standard Data Products* [Online]. Available: <http://aviris.jpl.nasa.gov/html/aviris.freedata.html>
- [43] R. N. Clark et al., "Imaging Spectroscopy: Earth and Planetary Remote Sensing With the Usgs Tetracorder and Expert Systems," *J. Geophys. Res.*, vol. 108, no. E12, pp. 5-1–5-44, Dec. 1993.
- [44] L. Miao and H. Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 3, pp. 765–777, Mar. 2007.
- [45] T.-H. Chan, C.-Y. Chi, Y.-M. Huang, and W.-K. Ma, "A convex analysis-based minimum-volume enclosing simplex algorithm for hyperspectral unmixing," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4418–4432, Nov. 2009.