



HAL
open science

Compressed sensing in mobile systems

Houria Haneche

► **To cite this version:**

Houria Haneche. Compressed sensing in mobile systems. Networking and Internet Architecture [cs.NI]. University of Science and Technology Houari Boumediene, 2020. English. NNT: . tel-03049353

HAL Id: tel-03049353

<https://hal.science/tel-03049353>

Submitted on 9 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research
University of Science and Technology Houari Boumediene
Faculty of Electronics and Computer Science



THESIS

Presented in total fulfillment of the requirements for the degree of
Doctor of Philosophy

In : Electronics

Speciality : Telecommunications and Information Processing

By: Houria HANECHÉ

TITLE

Compressed sensing in mobile systems

Defended publicly on, November 26, 2020 before the jury composed of:

Mrs. Amina SERIR	Professor	USTHB	President
Mr. Bachir BOUDRAA	Professor	USTHB	Thesis supervisor
Mr. Abdeldjalil OUAHABI	Professor	Polytech Tours	Thesis co-supervisor
Mr. Adel BELOUHRANI	Professor	ANPA	Examiner
Mrs. Fatma Zohra CHELALI	Associate Professor	USTHB	Examiner
Mr. Abdelaziz OUAMRI	Professor	USTO	Invited

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université des Sciences et de la Technologie Houari Boumediene
Faculté d'Electronique et d'Informatique



THÈSE DE DOCTORAT

Présentée pour l'obtention du grade de Docteur

En : Electronique

Spécialité : Télécommunication et Traitement de l'Information

Par: Houria HANECHÉ

THÈME

**Etude du paradigme acquisition compressée
dans les systèmes mobiles**

Soutenue publiquement le, Jeudi 26 Novembre 2020, devant le jury composé de:

M ^{me} Amina SERIR	Professeur	à l'USTHB	Présidente
M. Bachir BOUDRAA	Professeur	à l'USTHB	Directeur de thèse
M. Abdeldjalil OUAHABI	Professeur	à Polytech Tours	Co-directeur de thèse
M. Adel BELOUHRANI	Professeur	à l'ENPA	Examineur
M ^{me} Fatma Zohra CHELALI	Maître de Conférences A	à l'USTHB	Examinatrice
M. Abdelaziz OUAMRI	Professeur	à l'USTO	Invité

“The gold coin can be lost, and one finds it after it has been lost, but no force in the world can destroy a minute, nor restore it if it goes.”

Malek Bennabi

Abstract

Compressed sensing in mobile systems

We are interested in investigating the compressed sensing paradigm in the context of mobile systems to deal with the logistical and computational challenges in future communications. Due to the increasing number of connected objects (Internet of Things), the data exchanged in turn grows exponentially: this is the era of big data. This adds a level of intelligence to devices enabling them to communicate real-time data without a human being involved, effectively merging the digital and physical worlds. Hence the need to propose efficient techniques for compressing data from a variety of sources. Firstly, we propose simpler speech codecs based on quantization of compressed sensing measurements. The results show that the proposed codecs can be promising alternatives to current speech codecs. Secondly, we focus on the communication system. The behaviour of compressed sensing-source coding within the transmission chain is studied when undergoing real mobile communication-conditions. More specifically, we design new end-to-end mobile communication schemes based on compressed sensing. The proposed designs incorporate the compressed sensing-based speech codec instead of sampling signals at Nyquist rate then using a complex speech codec. Additionally, efficient techniques are chosen for channel compensation. The proposed systems show a simplified design, and allow reducing bit rate and processing load compared to actual communication systems based on adaptive multi-rate wideband (AMR-WB) speech codecs. The recovered speech has good quality and fair intelligibility scores when dramatic communication conditions are experienced (Rayleigh environment). In addition to reducing the computational burden for all the transmission steps, compressed sensing allows secure communications without additional costs. Thirdly, we consider the background noise coming from the environment. We propose a new speech enhancement method based on compressed sensing. In this approach we perform noise subtraction in the measurement domain before sparse recovery. Significant results are obtained showing that the proposed method is a good alternative to classical as well as prior compressed sensing-based speech enhancement methods, especially at low signal-to-noise ratios.

Keywords: Compressed sensing, mobile communication systems, speech codecs, speech enhancement

جامعة هوارى بومدين للعلوم والتكنولوجيا

ملخص

كلية الإلكترونيك و الإعلام الآلي

قسم الإتصالات السلكية واللاسلكية

متطلب تكميلي لنيل درجة الدكتوراه

دراسة نموذج الإستشعار المضغوط في أنظمة الاتصالات المتقلة

إعداد حورية حناش

نحن مهتمون باستغلال الاستشعار المضغوط في سياق الأنظمة المتقلة لمواجهة التحديات اللوجستية والحسابية في الاتصالات المستقبلية. أولاً ، نقترح طرقاً أبسط لضغط الكلام استناداً إلى القياس الكمي للقياسات التي تم جمعها بواسطة الاستشعار المضغوط. تظهر النتائج أن الطرق المقترحة يمكن أن تكون بدائل واعدة لتقنيات الترميز الصوتي الحالية. ثانياً ، نحن نركز على نظام الاتصالات. تتم دراسة سلوك ترميز المنبع المستند على الاستشعار المضغوط داخل سلسلة نظام الاتصالات في ظل ظروف الاتصالات المتقلة الحقيقية. على وجه التحديد، نحن نقوم بتصميم نظام اتصالات متقل جديد يعتمد على الاستشعار المضغوط. في التصميم المقترح يتم دمج ضغط الكلام المستند على الاستشعار المضغوط بدلاً من استشعار إشارات الكلام بمعدل نيكويست ثم استخدام تقنيات ترميز معقدة. بالإضافة إلى ذلك، يتم اختيار التقنيات الفعالة لمقاومة تأثيرات قناة الإرسال. النظام المقترح لديه تصميم مبسط وهو يقلل من معدلات الترميز وحمل المعالجة مقارنةً بأنظمة الاتصالات الحالية القائمة على تقنية الترميز المتكيف بمعدلات نقل متعددة عريض النطاق AMR-WB. يتمتع الكلام المسترجع بكفاءة جيدة من حيث الجودة و مقبولة من حيث الوضوح عندما تكون ظروف الاتصال سيئة (بيئة رايلي). بالإضافة إلى تقليل العبء الحسابي في جميع مراحل نقل الإشارة ، فإن الاستشعار المضغوط يتيح اتصالات آمنة دون تكاليف إضافية. ثالثاً ، نهتم بالضوضاء الخلفية الآتية من البيئة المحيطة. نقترح طريقة أخرى لتحسين الكلام تعتمد على الاستشعار المضغوط. في نهجنا ، نحن نقوم بطرح الضوضاء في مجال القياس قبل تطبيق خوارزمية الاسترجاع. تظهر النتائج الواعدة أن الطريقة المقترحة تعد بديلاً جيداً للطرق التقليدية وكذلك لطرق تحسين الكلام القائمة على الاستشعار المضغوط، خاصةً بالنسبة لنسب الإشارة إلى الضجيج الضئيلة.

كلمات مفتاحية: الاستشعار المضغوط، أنظمة الاتصالات المتقلة، ترميز الكلام، تحسين الكلام

Résumé

Etude du paradigme acquisition compressée dans les systèmes mobiles

Nous nous sommes intéressés par l'exploitation de l'acquisition compressée dans le contexte des systèmes mobiles pour faire face aux défis logistiques et calculatoires dans les communications futures. En raison du nombre croissant d'objets connectés (*Internet of Things*), les données échangées croissent à leur tour de façon exponentielle : c'est l'ère du *big data*. Cela ajoute une couche d'intelligence à ces objets leur permettant de communiquer des données en temps réel sans qu'un être humain soit impliqué, fusionnant ainsi efficacement les mondes numérique et physique. D'où la nécessité de proposer des techniques efficaces de compression de données provenant de sources diverses. Tout d'abord, nous proposons des codecs vocaux plus simples basés sur la quantification des mesures détectées par l'acquisition compressée. Les résultats montrent qu'ils peuvent être des alternatives prometteuses aux codecs vocaux actuels. Deuxièmement, nous nous concentrons sur le système de communication. Le comportement du codage source basé sur l'acquisition compressée à l'intérieur de la chaîne de transmission est étudié dans des conditions réelles de communication mobile. Plus précisément, nous concevons de nouveaux systèmes de communication mobile de bout en bout basés sur l'acquisition compressée. Les conceptions proposées intègrent des codecs vocaux basés sur l'acquisition compressée au lieu d'acquérir des signaux au débit de Nyquist puis utiliser un codec vocal complexe. De plus, des techniques efficaces sont choisies pour la compensation de canal de transmission. Les systèmes proposés présentent une conception simplifiée et permettent de réduire les débits binaires et la charge de traitement par rapport aux systèmes de communication actuels basés sur des codecs vocaux adaptatifs à large bande et débits multiples. La parole récupérée a une bonne qualité et une intelligibilité acceptable lorsque les conditions de communication sont dramatiques (environnement de Rayleigh). En plus de réduire la charge de calcul pour toutes les étapes de transmission, l'acquisition compressée permet des communications sécurisées sans coûts supplémentaires. Troisièmement, nous considérons le bruit de fond provenant de l'environnement. Nous proposons une autre méthode d'amélioration de la parole bruitée basée sur l'acquisition compressée. Dans notre approche, nous effectuons la soustraction de bruit dans le domaine de mesure avant d'appliquer l'algorithme de reconstruction. Les résultats prometteurs obtenus montrent que la méthode proposée est une bonne alternative aux méthodes classiques ainsi qu'aux méthodes d'amélioration de la parole basées sur l'acquisition compressée, surtout pour les faibles rapports signal-sur-bruit.

Mots clés: Acquisition compressée, systèmes de communications mobiles, codage de la parole, débruitage de la parole

Acknowledgements

I would like to express my sincere gratitude to my supervisor Prof. Bachir BOUDRAA and co-supervisor Prof. Abdeldjalil OUAHABI for their guidance, support, and encouragement, especially in the hard time of my research. I am also thankful to them for their patience and motivation.

I am also grateful to my committee members Dr. Amina SERIR, Dr. Adel BELOUHRANI, Dr. Abdelaziz OUAMRI, and Dr. Fatma Zohra CHELALI for the time and attention they devoted to my work. I would also like to thank them for their suggestions to improve the quality of this thesis.

This work is performed in the spoken communication and signal processing laboratory, I would like to thank all of my colleagues in the LCPTS research team.

I would like to thank my family for their continuous support and encouragement. I also owe my deepest appreciation to my grandmother who always prayed for me in my hard times.

In order not to forget anyone, my sincere thanks to all those who helped me to realize this dissertation.

Contents

Abstract	v
Acknowledgements	xi
List of Figures	xvii
List of Tables	xxiii
List of Abbreviations	xxv
Notations	xxvii
1 Introduction	1
1.1 Preface	1
1.2 Topic and Context	2
1.3 Research Focus and Scope	3
1.4 Value of the Research	4
1.5 Objectives of the Research	5
1.6 Thesis Outline	6
1.7 Publications	6
2 Compressed Sensing	9
2.1 Introduction	9
2.2 Brief History	9
2.3 Compressed Sensing Basics	11
2.3.1 Sparsity	12
Strict Sparsity and Compressibility	13
Sparsity Measures	14
How to Sparsify Signals?	17
2.3.2 Acquisition	18
Incoherence	18
Restricted Isometry Property	19

2.3.3	Signal Reconstruction	20
	Convex Optimization Approach	22
	Greedy Approach	23
	Thresholding Approach	27
	Combinatorial Approach	28
	Non-Convex Approach	29
	Bayesian Approach	30
2.4	Applications of Compressed Sensing	30
2.4.1	Compressive Imaging	30
	Single Pixel Camera	30
	Medical Imaging	32
	Seismic Imaging	32
	Radar Imaging	32
2.4.2	Biomedical Applications	33
2.4.3	Radar	33
2.4.4	Pattern Recognition	33
2.4.5	Video Processing	34
2.4.6	Manifolds Processing	34
2.4.7	Micro and Nano-Electronics and VLSI Applications	34
2.4.8	Analog-To-Information Conversion	34
2.4.9	Speech and Sound Processing	35
	Speech Coding	35
	Speech Enhancement	35
2.5	Compressed Sensing in Communication Systems	37
2.5.1	Sparse Channel Estimation	37
2.5.2	Wireless Sensor Networks	38
2.5.3	Ultra Wideband Systems	38
2.5.4	Cognitive Radio	38
2.5.5	Array Signal Processing	39
2.5.6	Multiple Access Scheme	40
2.5.7	Network Tomography	40
2.5.8	Multimedia Coding and Communication	41
2.5.9	Information Security	42
2.6	Conclusion	42
3	Compressed Sensing-Based Speech Codec	43
3.1	Introduction	43
3.2	Motivation	43

3.3	Overview of Quantization Methods	45
3.3.1	Scalar Quantization	45
3.3.2	Vector Quantization	45
3.4	Speech Codec Based on Multistage Vector Quantization of Compressed Sensing Measurements	47
3.5	Speech Codec Based on Split Vector Quantization of Compressed Sensing Measurements	48
3.6	Speech Codec Based on Split-Multistage Vector Quantization of Compressed Sensing Measurements	49
3.7	Experiments	51
3.7.1	Simulation Setup	53
3.7.2	Performance Measures	53
3.7.3	Results and Discussion	54
	Evaluation of CS-Codec using Scalar Quantization	54
	Evaluation of CS-Codex using Vector Quantizations	56
	Comparison with Other Compressed Sensing-Based Codex	74
	Comparison with currently used compression standards	75
3.8	Conclusion	76
4	Designing a New Mobile Communication System by Exploiting Compressed Sensing	79
4.1	Introduction	79
4.2	Motivation	79
4.3	The Proposed End-to-End Communication System Model	81
4.3.1	Compressed Sensing-Based Source Coding	81
4.3.2	Fading and Channel Modelling	82
	Multipath fading effects in mobile communication systems	82
	Modelling the propagation channel	83
4.3.3	Multipath Fading Compensation	84
	Anti-Multipath Systems	84
	Anti-Fading Systems	85
4.3.4	Channel Coding	87
4.4	Improving the System for Lower Bitrates	88
4.5	Experiments	88
4.5.1	Simulation Setup and Performance Measures	95
4.5.2	Results	96
	Evaluation of the Proposed Communication System	96

	Evaluation of the improved version of the Proposed Communication System	101
4.5.3	Discussion	102
	Quality and Intelligibility Assessment	104
	About Complexity and Security Aspects	107
4.6	Conclusion	108
5	Speech Enhancement Using Compressed Sensing	111
5.1	Introduction	111
5.2	Motivation	111
5.3	Brief Description of Some Conventional Speech Enhancement Methods	113
5.3.1	Spectral Subtraction	113
5.3.2	Wiener Filtering	114
5.3.3	Wavelet Denoising	115
5.4	Compressed Sensing-Based Speech Enhancement	116
5.5	Proposed compressed sensing-based speech enhancement method	116
5.5.1	System Overview	117
5.5.2	Mathematical Formulation of the Proposed Speech enhancement process	118
5.6	Experiments	121
5.6.1	Experimental Setting	122
5.6.2	Performance Measures	123
5.6.3	Results and Discussion	123
5.7	Conclusion	134
6	Conclusions and Perspectives	135
	Index	139
	Bibliography	141

List of Figures

2.1	Classical sampling then compression versus compressed sampling schemes. <i>Note that compressed sampling performs sampling and compression simultaneously.</i>	11
2.2	Compressed sensing scheme. <i>The speech signal \mathbf{s} of length N is sparsified by projecting it on the representation basis Ψ that can be Fourier basis, wavelets, ... As a result, we obtain the sparse vector θ (of length N) that contains only K non null values, so we say that it is K-sparse. The acquisition is performed by multiplying the measurement matrix Φ of length $M \times N$ by the signal \mathbf{s}. Hence, we obtain the acquired signal \mathbf{y} (measurement vector) that contains only M measurements, such that $K < M \ll N$</i>	12
2.3	Serial greedy algorithms steps	25
2.4	Combinatorial approach (using count min/median) steps	29
3.1	The proposed compressed sensing-based speech codec using split vector quantization (SVQ). <i>(This codec performs SVQ quantization of the observation vector \mathbf{y}. Note that ω is the measurement noise due to compressed sensing.)</i>	44
3.2	The proposed compressed sensing-based speech codec using split-multistage vector quantization (S-MSVQ). <i>(This codec performs S-MSVQ quantization of the observation vector \mathbf{y}. Note that ω is the measurement noise due to compressed sensing.)</i>	44
3.3	Diagram of scalar quantization	46
3.4	Diagram of vector quantization	46
3.5	Diagram of three stages MSVQ <i>(note that at stage t ($1 \leq t \leq 3$) the vector of indexes is $\mathbf{I}^t = [i_1^t i_2^t \dots]$).</i>	47
3.6	Diagram of three splits vector quantizer <i>(note that for a split p ($1 \leq p \leq 3$) the vector of indexes is $\mathbf{I}_p = [ip_1 ip_2 \dots]$).</i>	48

3.7	Codebooks Generation for three-parts three-stages split-multistage vector quantizer (S-MSVQ). <i>The training sequence is partitioned to three splits. LBG algorithm is performed on the three splits at the 1st stage, and on the quantization error (of the previous stage) for the 2nd and the 3rd stages. As a result, nine sub-codebooks are created</i>	50
3.8	Diagram of three-splits three-stages split-multistage vector quantizer (S-MSVQ). <i>The quantized observation vector \hat{y} is obtained by summing the quantized subvectors of the three stages and concatenating the resultant splits</i>	51
3.9	Simulation system.	52
3.10	The PESQ and SII scales	54
3.11	Performance of speech compression through scalar quantization of CS measurements for an <i>nbps</i> varying from 1 to 10 and for various compression ratios (<i>C</i>).	55
3.12	The effect of scalar quantization on compressed sensing recovery. <i>Note that the performances of speech codec based on compressed sensing only (without quantization) and those of the speech codec based on scalar quantization of CS measurements (with $nbps = 4$) are similar for various compression ratios <i>C</i>.</i>	57
3.13	Performance of speech compression through vector quantization of compressed sensing measurements. <i>Note that the bitrate is fixed to 1.6 kbit/s the compression ratio (<i>C</i>) is varying from 10% to 90%.</i>	59
3.14	Performance of speech compression through split vector quantization of compressed sensing measurements. <i>Note that the bitrate is fixed to 1.6 kbit/s the compression ratio (<i>C</i>) is varying from 10% to 90%.</i>	60
3.15	Performance of speech compression through multistage vector quantization of compressed sensing measurements. <i>Note that the bitrate is fixed to 1.6 kbit/s the compression ratio (<i>C</i>) is varying from 10% to 90%.</i>	62
3.16	Performance of speech compression through split-multistage vector quantization of compressed sensing measurements. <i>Note that the bitrate is fixed to 1.6 kbit/s the compression ratio (<i>C</i>) is varying from 10% to 90%.</i>	64
3.17	Performance of speech compression through vector quantization of compressed sensing measurements. <i>Note that the compression ratio (<i>C</i>) is fixed to 10% and the number of bits per vector is varying from 10 to 18 bits/vector.</i>	67
3.18	Performance of speech compression through split vector quantization of compressed sensing measurements. <i>Note that the compression ratio (<i>C</i>) is fixed to 10%, and the number of bits per vector is varying from 10 to 18 bits/vector, which is equivalent to $1.2 \leq BR \leq 2.16$ kbit/s.</i>	69

3.19	Performance of speech compression through multistage vector quantization of compressed sensing measurements. <i>Note that the compression ratio (C) is fixed to 10% and the number of bits per vector is varying from 10 to 18 bits/vector.</i>	70
3.20	Performance of speech compression through split-multistage vector quantization of compressed sensing measurements. <i>Note that the compression ratio (C) is fixed to 10% and the number of bits per vector is varying from 10 to 18 bits/vector.</i>	72
4.1	General diagram of communication systems	80
4.2	General diagram of the proposed communication scheme (<i>note that this system integrates compressed sensing-speech coding instead of the current codecs standards such as G.722.2</i>)	81
4.3	Proposed compressed sensing-based speech codec	82
4.4	Proposed compressed sensing-based speech codec	88
4.5	Mobile communication system using compressed sensing.	89
4.6	Flowchart of compressed sensing-source coding.	90
4.7	An extract of 1000 samples of original speech signal (top), and its sparse representation (bottom). (<i>Observe that the wavelet representation is sparser than the temporal one</i>).	97
4.8	Performance of selection combining diversity with OFDM modulation in Rayleigh channel. (<i>Observe that with SC the BERs are reduced when the number of receiving antennas increases: Around 8 dB improvement at 10^{-2} BER point with $n_{Rx} = 2$ and 10 dB with $n_{Rx} = 3$</i>).	97
4.9	Performance of equal gain diversity combining with OFDM modulation in Rayleigh channel. (<i>Observe that with EGC the BERs are reduced when the number of receiving antennas increases: Around 9 dB improvement at 10^{-2} BER point with $n_{Rx} = 2$ and 12 dB with $n_{Rx} = 3$</i>).	98
4.10	Performance of maximum ratio diversity combining with OFDM modulation in Rayleigh channel. (<i>Observe that with MRC the BERs are reduced when the number of receiving antennas increases: Around 10 dB improvement at 10^{-2} BER point with $n_{Rx} = 2$ and 13 dB with $n_{Rx} = 3$</i>).	98
4.11	Performance comparison of diversity combining techniques with OFDM modulation in Rayleigh channel ($n_{Rx}=3$). (<i>Observe that the BERs of MRC are the smallest</i>).	99

4.12	Constellation diagrams of the transmitted signal (in red asterisks) and the received signal (in blue dots) after equalization with receive diversity ($BR = 12.8$ kbit/s; $n_{Rx}=2$; MRC; $E_b/N_o= 20$ dB). (<i>Observe that the QAM-16 constellation is recovered with few distortion due to the white Gaussian noise.</i>)	100
4.13	The original (top) and the recovered (bottom) speech signals. (<i>Note that the mean-squared error of this example is 1.2×10^{-3}, and the PESQ is 3.43.</i>)	100
4.14	PESQ performances of the reconstructed signal for the proposed communication system, $BR= 8.85$ kbit/s. (<i>Observe that the best quality is obtained when using 2-stages 3-splits S-MSVQ for a $CR=20\%$</i>)	102
4.15	Intelligibility performances of reconstructed signal for the proposed communication system, $BR= 8.85$ kbit/s. (<i>Observe that the intelligibility score increases with the rising of CR</i>)	104
5.1	Block Diagram of the proposed CS-based speech enhancement method. <i>Note that for each frame, a decision speech/silence is performed and a processing is done depending on the result.</i>	117
5.2	The STOI scale.	123
5.3	Segmental SNR improvements of the proposed CS-based speech enhancement method, Wang16 and CoSaMPSE in various noise conditions : (a) <i>babble</i> , (b) <i>police siren</i> , (c) <i>piano</i> , and (d) <i>market</i> . The segmental SNR improvements of Wang16 and CoSaMPSE are not available for the input SNRs -10 dB and -5 dB.	127
5.4	PESQ performances of the proposed CS-based speech enhancement method, Wang16 and CoSaMPSE in various noise conditions: (a) <i>babble</i> , (b) <i>police siren</i> , (c) <i>piano</i> , and (d) <i>market</i> . The PESQ scores of Wang16 and CoSaMPSE are not available for the input SNRs -10 dB and -5 dB.	128
5.5	STOI performances of the proposed CS-based speech enhancement method, Wang16 and CoSaMPSE in various noise conditions: (a) <i>babble</i> , (b) <i>police siren</i> , (c) <i>piano</i> , and (d) <i>market</i> . The STOI values of Wang16 and CoSaMPSE are not available for the input SNRs -10 dB and -5 dB.	129
5.6	PESQ performances of the proposed CS-based speech enhancement method and supervised NMF in <i>white</i> noise.	130
5.7	PESQ performances of the proposed CS-based speech enhancement method, supervised NMF, CJSR, GDL, and RNG-WWF in <i>factory</i> noise.	130
5.8	PESQ performances of the proposed CS-based speech enhancement method, RNG-WWF, and GDL in <i>babble</i> noise.	131

- 5.9 Clean (top), noisy (middle), and enhanced (bottom) speech waveforms
. Note that This signals is the sentence “Wipe the grease off his dirty face” (
|'waɪp ðə 'ɡri:s ɔ:f ɪz 'dɜ:ti 'feɪs|) pronounced by a male. The noisy signal
is obtained by corrupting the clean speech by babble noise at 0 dB. 132

List of Tables

2.1	Compressed sensing reconstruction approaches	31
3.1	The average performance scores of the recovered speech signals after compression ($BR = 1.6$ kbit/s) using different quantization techniques (namely, scalar quantization, vector quantization, three-splits split vector quantization, two-stages multistage vector quantization, and two-stages three-splits split-multistage vector quantization) with CS measurements. <i>Note that the bitrate is fixed to 1.6 kbit/s the compression ratio (C) is varying from 10% to 90%. The best scores for each quantization method are shown in bold.</i>	65
3.2	The average performance scores of the recovered speech signals after compression using different quantization techniques (namely, scalar quantization, vector quantization, three-splits vector quantization, two-stages vector quantization, and two-stages three-splits vector quantization) with compressed sensing measurements. <i>Note that the compression ratio (C) is fixed to 10% and the number of bits per vector are varying from 10 to 18 bits.</i>	73
3.3	Performance measures of the two proposed speech codecs (compressed sensing with three-splits split vector quantization and compressed sensing with two-stages three-splits split-multistage vector quantization)	75
3.4	PESQ comparison of the proposed CS-codec with AMR and AMR-WB in noise free and additive white Gaussian noise environment.	75
4.1	Simulation parameters	96
4.2	The average PESQ scores of the recovered speech signals ($nbps = 4$, $E_b/N_o = 10$ dB).	101
4.3	The average SII values of the recovered speech signals ($nbps = 4$, $E_b/N_o = 10$ dB).	101
4.4	Performance of the proposed communication system and its improved version for $BR = 8.85$ kbit/s	103

5.1	PESQ performances of the proposed CS-based speech enhancement method for different sparsity values of K-SVD in 10 dB <i>babble</i> noise. .	122
5.2	SNRseg improvement, PESQ, and STOI performances of the CS-based speech enhancement method (without VAD), the proposed CS-based speech enhancement method, and the proposed CS-based speech enhancement method with the Oracle mask in <i>babble</i> noise.	124
5.3	Methods used for comparison	126
5.4	Segmental SNR improvements of the proposed CS-based speech enhancement method, Low13, Wang16, and CoSaMPSE in <i>babble</i> noise.	132

List of Abbreviations

5G	5 Generation
ADC	Analog-to-Digital Converter
AIC	Analog-to-Information Conversion
AMR-WB	Adaptive Multi-Rate Wideband
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
CDMA	Code-Division Multiple-Access
CJSR	Complementary Joint Sparse Representations
CoSaMP	Compressive Sampling Matching Pursuit
CoSaMPSE	Compressive Sampling Matching Pursuit Speech Enhancement
CP	Cyclic Prefix
CS	Compressed Sensing
DMD	Digital Micro-mirror Devices
DPCM	Differential Pulse Code Modulation
ECG	ElectroCardioGram
EEG	ElectroEncephaloGraphic
EGC	Equal Gain Combining
FOCUSS	FOCal Underdetermined System Solution
GPR	Ground Penetrating Radar
GDL	Generative Dictionary Learning
IHT	Iterative Hard Thresholding
IoT	Internet of Things
IRLS	Iteratively Reweighted Least Squares
ISAR	Inverse Synthetic Aperture Radar
ISI	InterSymbol Interference
K-SVD	K-Singular Value Decomposition
LASSO	Least Absolute Shrinkage and Selection Operator
LBG	Linde, Buzo and Gray
LPC	Linear Predictive Coding
MIMO	Multiple Input Multiple Output
MISO	Multiple Input Single Output)

MRC	Maximum Ratio Combining
MRI	Magnetic Resonance Imaging
MSE	Mean-Squared Error
MSVQ	MultiStage Vector Quantization
OFDM	Orthogonal Frequency-Division Multiplexing
OMP	Orthogonal Matching Pursuit
PESQ	Perceptual Evaluation of Speech Quality
QAM	Quadrature Amplitude Modulation
S-MSVQ	Split-MultiStage Vector Quantization
SAR	Synthetic Aperture Radar
SC	Selection Combining
SII	Speech Intelligibility Index
SIMO	Single Input Multiple Output
SNR	Signal to Noise Ratio
SNRseg	Segmental Signal to Noise Ratio
SQ	Scalar Quantization
STFT	Short-Time Fourier Transform
STOI	Short-Time Objective Intelligibility
SVQ	Split Vector Quantization
TWR	Through the Wall imaging Radar
VAD	Voice Activity Detection
VLSI	Very Large Scale Integration
VoLTE	Voice over Long-Term Evolution
VQ	Vector Quantization

Notations

Boldface capital letters	matrices
Boldface small letters	vectors
Non-bold letters	scalars
$[\cdot]^T$	transpose
$\langle \cdot, \cdot \rangle$	scalar product
$\text{card}(\cdot)$	cardinality
$ \cdot $	absolute value
$\ \cdot\ _p$	ℓ_p norm
$\tanh(\cdot)$	hyperbolic tangent function
$\log(\cdot)$	logarithm function
$\exp(\cdot)$	exponential function
*	convolution operator
θ_i	i th element of vector θ
Θ^f	f th frame of matrix Θ
$\theta^{(i)}$	θ at iteration i
\mathbb{R}	the set of real numbers
\mathbb{C}	the set of complex numbers

*Dedicated to those
who gave me the force
to success ...*

Chapter 1

Introduction

1.1 Preface

In 1948, Claude Elwood Shannon published his well-known sampling theorem (Shannon, 1948). One year later, he proposed a novel model of communication system (Shannon, 1949), setting the foundation of information theory. It is one of the theoretical works that have had the greatest impact on modern electrical engineering (Unser, 2000). His pioneering work and the work of Kotelnikov (Kotel'nikov, 1933), Nyquist (Nyquist, 1928), and Whittaker (Whittaker, 1915) form the theoretical foundation on sampling continuous-time band-limited signals. Sampling theorem plays a capital role in signal processing and communications (Femmam, M'Sirdi, and Ouahabi, 2001). It shows us how to convert an analogue signal into a sequence of numbers, which can be processed by digital systems. Their results say that an exact recovery can be achieved from a set of uniformly spaced samples taken at twice the highest frequency present in the signal of interest (the so-called Nyquist rate). Capitalizing on this discovery, much of signal processing has moved from the analogue to the digital domain. This has allowed designing processing systems that are more robust, flexible, cheaper, and consequently, more widely used than their analogue counterparts. This success resulted in growing the amount of generated data from a trickle to a torrent. To overcome some practical limitations, primarily the constraint regarding the sampling rate, alternative sampling protocols have been proposed over the years, but without significant variations in the theoretical framework (Unser, 2000). Even after more than 70 years, signal acquisition systems are based on Shannon's sampling theorem and the corresponding reconstruction formula.

In 2006, a new sampling paradigm was defined by two major papers; namely, (Candès, Romberg, and Tao, 2006) and (Donoho, 2006). It says that an exact recovery of signal with sparse representation can be achieved from a small set of linear, non-adaptive measurements. This allows sensing sparse signals by collecting far

fewer measurements than required by Shannon; hence the name “compressed sensing”. Compressed sensing (CS) has already drawn an extensive interest in various domains. For example, in medical imaging (Yang, Qin, and Wu, 2019; Mardani et al., 2019), where it permits an acceleration in magnetic resonance imaging (MRI) while preserving diagnostic quality. Other applications are, especially in communication systems (Zhang et al., 2019; Qian, Fu, and Sidiropoulos, 2019; He et al., 2019), digital signal processing for analog-to-digital conversion (Pelissier and Studer, 2018; Xu, Zhang, and Kim, 2019), reconstruction algorithms design (Andráš et al., 2018), *etc.*

1.2 Topic and Context

Mobile technology was a mystery two decades ago but now, it is taking center stage. Since the arrival of the mobile, it has helped humans in many ways. In addition of making calls and sending messages, variety of services are offered, for example, making purchases, video-conferencing, gaming, *etc.* Mobile systems hold a lot more features in the future to meet even the most of our basic needs and to make life a lot easier. Advances in mobile systems have resulted not only in improving the users connectivity, but also in the emergence of devices connected over the internet, known as the internet of things (IoT). Although still at a nascent stage, there are already some 19 million connected devices in Algeria (La Rédaction, 2018).

Future mobile systems plan to provide high data rates of more than 100 Gbps. In the other side, users demand for high-quality multimedia, high speed, high reliability, and real-time communication at any time and anywhere.

In view of this, challenges arise in this domain. Indeed, current communication systems allow high bandwidth mobile service, but it seems that they will not be able to meet the emerging user’s expectations for the future mobile communications, where billions of devices require secured Gigabit wireless connectivity (Internet of Things). To be more specific, the extensive growth of applications and the fact that we enter the era of big data lead to a continuous explosion in the amount of acquired samples from sensing systems. This causes an increasing burden at every stage of the communication pipeline, from the initial acquisition of the data to the subsequent transmission, storage, and analysis. What’s more, inefficiencies in spectrum usage and physical resources and a very high energy consumption rise, and this causes serious logistic problems. Another key thing to remember is that current systems depend on complex compression techniques, which aim to find a more concise signal representation with less distortion, to address the issue of huge data amount, and this

increases the computational load. Furthermore, lacks in robustness and reliability as well as security threats increase (Wunder et al., 2015; Gao et al., 2018).

To address these drawbacks, researches on new technologies for future communication systems are in progress. Providing higher rates is a purpose, but they are more concerned by reducing the processing load and the memory requirements, and enhancing the battery life and the security. All things considered, our topic is to investigate the new paradigm “compressed sensing” in mobile systems. Notably, we propose low-cost solutions to the aforementioned shortcomings based on sparse signal processing.

1.3 Research Focus and Scope

Not surprisingly, voice communication is the most important form of telecommunications as it maintains good relationships, and save time and money. It played and continues to play a significant role in many applications, for instance, in business for telemarketing, teleconferencing, and telecommuting (also known as working from home).

The consumption of physical resources used for mobile communications depends on the circulating data volume. The latter knows an exponential growth due to the high and continuous demand for the use of multimedia, including speech. Therefore, there is a serious need to find better methods for compression and communication. As such, our research focus is speech coding and communication.

To date, current communication systems perform speech coding after acquiring signals using the classical Shannon-Nyquist sampling theorem. The performance of systems is, thereby, limited because such acquisition becomes a burden, when only few coefficients are needed for its representation. Naturally, this may be unnecessarily wasteful, i.e., it can require expensive hardware, consume valuable power, *etc.* This problem worsens for high band signals. For instance, broadband telephony gradually replaces conventional telephony (which allows voice transmission between 300 and 3400 Hz), as is the case for voice over long-term evolution (VoLTE). For such systems wideband speech, covering frequencies in the range of 50 – 7000 Hz; thus allowing audio with richer tones and better quality, is transmitted. A sampling frequency of 16 kHz is required; hence the high load, memory, and energy requirements. Additionally, current speech coding standards, such as ITU-T G.722.1, adaptive multi-rate wideband (AMR-WB, also normalized as ITU-T G.722.2),... require high computational complexity, and this may increase the network load.

In this thesis, we exploit the dimensionality reduction property of CS to design new methods for speech coding. In addition, we propose new end-to-end communication schemes based on CS (i.e., the proposed compressed sensing-based speech codec and its improved version).

Furthermore, we focus on the speech enhancement as the latter is an important task in many areas including speech coding. We propose a new speech enhancement method based on CS to improve the quality of speech corrupted by background noise, coming from the environment such as in restaurants, market, and factory.

1.4 Value of the Research

In many important applications, building devices to acquire samples at the necessary Nyquist rate may be too costly. Moreover, using compression to deal with high dimensional data may involve computational load to the system. To address the logistical and computational challenges, we perform simultaneously the acquisition and the compression using CS. In other words, we incorporate dimensionality reduction into the sensing process itself by collecting fewer samples than suggested by the classical Shannon-Nyquist sampling theorem. Since the emerging of CS, it has been most widely applied in the image processing area where it has brought great advances in compression. While there has been previous studies in the speech processing area, the later lacks investigations on the application of CS for speech coding. Only few works have focused on speech compression using CS. In (Giacobello et al., 2012), the authors have exploited CS to improve the linear prediction coding (LPC) performance. (Al-Azawi and Gaze, 2018) has introduced a combined compression/encryption method for speech using CS. Ji *et al.* in (Ji, Zhu, and Champagne, 2019) have proposed a dictionary learning technique based on the recurrent neural network for compressive speech sensing. They have also proposed a new technique for the extraction of LPC coefficients. In this dissertation, we propose simpler speech codecs based on quantization of CS measurements. The results shows that they can be promising alternatives to current speech codecs.

Moreover, no study, as far as we know, has considered the communication system. In this thesis, we demonstrate an application of CS for speech compression in the context of mobile systems. More specifically, we design new end-to-end mobile communication schemes based on CS. The proposed designs use a CS-based speech codec instead of acquiring signals at Nyquist rate then using complex speech codecs. Additionally, efficient techniques are chosen for channel compensation. In other words, the behaviour of CS-source coding within the transmission chain is studied

when undergoing real mobile communication-conditions. The proposed systems show a simplified design, and allows reducing bit rates and processing load compared to actual communication systems based on AMR-WB speech codecs. The recovered speech has good quality and fair intelligibility scores when dramatic communication conditions are experienced (Rayleigh environment). In addition to reducing the processing load in all the transmission steps, CS allows secure communications without additional costs.

For speech enhancement, few studies have used CS (Sigg, Dikk, and Buhmann, 2012; Jančovič, Zou, and Köküer, 2012; Low, Pham, and Venkatesh, 2013; Wu, Zhu, and Swamy, 2013; Wu, Zhu, and Swamy, 2014; Ramdas, Mishra, and Gorthi, 2015; Luo et al., 2016; Wang et al., 2016). They are mainly based on sparse recovery of speech from a mixture of clean speech and noise. As the CS reconstruction quality degrades at low signal-to-noise ratios (SNRs), the performance of these methods deteriorates at low SNRs. In our thesis, we address this problem by performing noise subtraction in the measurement domain before sparse recovery. Promising results are obtained showing that the proposed method is a good alternative to classical and CS-based speech enhancement methods.

1.5 Objectives of the Research

The aim of this thesis is to contribute to the development of a mobile system for speech communication by exploiting CS. In this perspective, the following objectives are underlined:

- Design a new speech codec based on CS by exploiting sparsity of speech in a chosen basis.
- Investigate the behaviour of the proposed codec in dramatic mobile communications.
- Propose a new end-to-end mobile communication scheme by incorporating the designed CS-based speech codec, and choosing the suitable techniques for channel compensation to meet the requirements of future mobile communications such as 5G systems.
- Propose a new speech enhancement method based on CS to denoise speech from background noise.

1.6 Thesis Outline

Chapter 2 provides the basic concepts of CS. The mathematical formulation is given, in addition to an overview of its applications in many areas, including those in communications.

Chapter 3 describes our first contribution. It consists of designing a new speech codec based on CS. Two codecs are proposed, and new results are presented for speech compression.

Chapter 4 develops our second contribution on designing a new end-to-end mobile communication scheme. The proposed system is described and results for speech coding and communication are provided. In addition, an improved version of the designed communication system for low bitrate is described.

Chapter 5 describes our third contribution in speech enhancement. It presents the new CS-based speech enhancement method.

Chapter 6 provides conclusions and future perspectives.

1.7 Publications

Journal Papers

Haneche, H., Ouahabi, A., Boudraa, B. (2019). New mobile communication system design for Rayleigh environments based on compressed sensing-source coding, *IET Communications* 13 (15). pp. 2375-2385. DOI: 10.1049/iet-com.2018.5348

Haneche, H., Boudraa, B., Ouahabi, A. (2020). A new way to enhance speech signal based on compressed sensing. *Measurement* 151 (1). DOI: 10.1016/j.measurement.2019.107117

Haneche, H., Ouahabi, A., Boudraa, B. Compressed Sensing-Speech Coding Scheme for Mobile Communications. *Circuits, Systems, and Signal Processing (Accepted)*

International Conferences

Haneche, H., Boudraa, B., Ouahabi, A. (2015, May). Application de la CS dans les systèmes mobiles. In *Colloque International TELECOM'2015 & 9èmes JFMMA* (pp. 1-5). May 13-15 2015, Meknes, Morocco.

Haneche, H., Boudraa, B., Ouahabi, A. (2018, October). Compressed Sensing Investigation in an End-to-End Rayleigh Communication System: Speech Compression.

In *2018 International Conference on Smart Communications in Network Technologies (SaCoNeT)* (pp. 73-77). IEEE. October 27-31 2018, Eloued, Algeria.

Haneche, H., Boudraa, B., Ouahabi, A. (2018, November). Split Vector Quantization of Compressive Sampling Measurements for Speech Compression. In *2018 International Conference on Signal, Image, Vision and their Applications (SIVA)* (pp. 1-6). IEEE. November 26-27 2018, Guelma, Algeria.

Haneche, H., Boudraa, B., Ouahabi, A. (2018, October). Speech Enhancement Using Compressed Sensing-based method. In *2018 International Conference on Electrical Sciences and Technologies in Maghreb (CISTEM)* (pp. 1-6). IEEE. October 28-31 2018, Algiers, Algeria.

National Conferences

Haneche, H., Boudraa, B., Ouahabi, A. (2015, Janvier). L'acquisition compressée dans les systèmes mobiles. In *2èmes Journées du Laboratoire de Communication Parlée et de Traitement des Signaux (JLCPTS 2015)* (pp. 1-4). January 14-15 2015, Algiers, Algeria.

Haneche, H., Boudraa, B., Ouahabi, A. (2016, December). An Optimized Design of Mobile Communication System. In *Workshop of the Speech Communication and Signal Processing Laboratory (JLCPTS 2016)* (pp. 1-3). December 5 2016, Algiers, Algeria.

Chapter 2

Compressed Sensing

2.1 Introduction

The Nyquist-Shannon sampling theorem, saying that the sampling rate must be at least twice the signal bandwidth, has been widely used for signals acquisition for over fifty years. Nowadays, in many applications (e.g., broadband telephony) the Nyquist rate is so high; hence an excessive number of samples is acquired. This entails a compression prior to storage or transmission, and makes the acquisition process costly. The compressed sensing (CS) paradigm is an efficient signal processing technique for simultaneously sensing and compressing signals. Compressed sensing stems from the idea that it is not necessary to invest a lot of resources to acquire a sparse signal, where most of its inputs are anyway null (generally, in a transform domain). In fact, it should be possible to collect only a few measures that allow future reconstruction rather than acquire the whole signal then compress it. Thanks to sparsity and under some constraints, CS allows signal recovery from fewer measures than required by Nyquist-Shannon sampling theorem. Since its initiation in 2006, it has gained an extensive interest in many fields such as electrical engineering, applied mathematics, computer science, and telecommunications. The purpose of this chapter is to provide the basic concepts of CS and an overview of its applications. It begins by giving a mathematical formulation for CS, explaining the sparsity, the sensing matrix, and the CS reconstruction. Next, it gives an overview of its applications in many areas including those in communications.

2.2 Brief History

The problems connected to CS can be found in earlier papers. Presumably, Prony's method was the first proposition that has link with CS (Foucart and Rauhut, 2013). This method, dated 1795, is proposed by Prony in (Prony, 1795) for estimating the

frequencies, the amplitudes, and the phases from a uniformly sampled signals by solving an eigenvalue problem (Potts and Tasche, 2010).

Compressed sensing draws on ℓ_1 minimization. The latter was proposed by Logan in 1965 for sparse frequency estimation (Logan, 1965). Another early work on ℓ_1 minimization is (Donoho and Logan, 1992). In seismology, geophysicists succeeded in constructing images of reflective layers within the earth from incomplete (non satisfactory) data using the ℓ_1 minimization. In 1970s, Claerbout and Muir gave attractive alternative of least square solutions (Claerbout and Muir, 1973). Kashin (Kašin, 1977) and Gluskin (Gluskin, 1984) gave norms for random matrices. They gave precise bounds on the performance of the recovery of sparse vectors from linear measurements. Santosa and Symes (Santosa and Symes, 1986) in mid eighties observed that ℓ_1 minimization can successfully used for the recovery of sparse spike trains. In 1990s, Rudin, Osher, and Fatemi (Rudin, Osher, and Fatemi, 1992) utilized total variation minimization, which is very close to ℓ_1 minimization, in image processing. The work of Tibshirani (Tibshirani, 1996) on the least absolute shrinkage and selection operator (LASSO) remarkably popularized the use of ℓ_1 minimization and related greedy methods in statistics.

In computer science, through the theory of sketching algorithms, sparsity was considered (before the apparition of CS) when recovering huge data sets (e.g., Internet data stream) from undersampled data at a sublinear time. In this case, only the locations and values of nonzero entries need to be reported. Group testing (Du and Hwang, 1993), which is closely related in spirit to CS (Atia and Saligrama, 2012), was used. Group testing was proposed by Dorfman during the second world war (Dorfman, 1943) to detect soldiers affected by syphilis using a reduced number of tests. Its goal is to efficiently recover a small number of items from a large number of entries while reducing the required number of tests (measurements).

It is important to notice that studies on the sparse recovery problem started in the 1990s with the following works (Chen, Donoho, and Saunders, 1998; Mallat and Zhang, 1993; Natarajan, 1995). Moreover, conditions permitting the recovery of the sparsest solution by greedy methods and ℓ_1 minimization were developed prior to the CS apparition, e.g., (Donoho and Huo, 2001; Fuchs, 2004; Gilbert, Muthukrishnan, and Strauss, 2003; Gribonval and Nielsen, 2003; Tropp, 2004; Tropp, 2006).

The idea of CS got a new life in 2004 with the results achieved by David Donoho, Emmanuel Candes, Justin Romberg and Terence Tao. When the two fundamental papers; namely, (Donoho, 2006) and (Candès, Romberg, and Tao, 2006) were published, the term “compressed sensing” was coined out and this field was initiated. These two

papers are the first ones who gave the mathematical foundation of CS. They proved the effectiveness of choosing random measurement matrix (which allows a minimal number of linear measurements) with ℓ_1 minimization in solving underdetermined systems of equations in various signal processing tasks.

Since then, a plethora of works on CS are already accomplished, and exploited for a variety of applications in astronomy, biology, medicine, radar, seismology, *etc.* It is worth noting that, generally, the terminology “compressed sensing” is used interchangeably with “sparse recovery”.

2.3 Compressed Sensing Basics

Compressed sensing (also known as compressive sensing, compressive sampling, or sparse sampling) was launched in 2006 by two pioneer papers; namely, (Donoho, 2006) by Donoho and (Candès, Romberg, and Tao, 2006) by Candes, Romberg, and Tao. It aims to reconstruct sparse signals from a small number of random projections. In fact, the idea is to acquire only a few measurements allowing future reconstruction (under some constraints) rather than acquiring a large number of samples at Nyquist rate, and then performing compression (see Fig. 2.1). This may reduce the resources

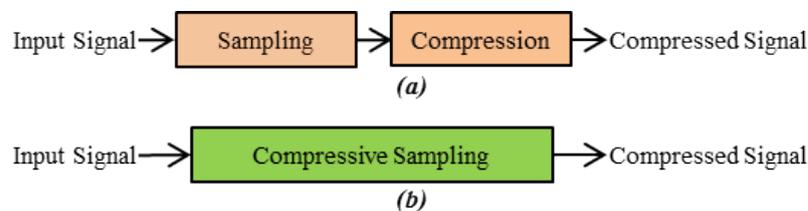


FIGURE 2.1: Classical sampling then compression versus compressed sampling schemes. *Note that compressed sampling performs sampling and compression simultaneously.*

needed for sampling signals whose most of their inputs are anyway null. The theoretical background of CS has connections to many other fields such as numerical linear algebra, optimization theory, random matrix theory, applied harmonic analysis, *etc.* The compressed sensing process is composed of three main parts as presented in Fig. 2.2:

1. sparse representation using a sparse basis (representation basis) to obtain a vector with only few non zero values;
2. acquisition using a sensing matrix to obtain the CS measurements; and
3. compressed sensing reconstruction using a sparse recovery algorithm to recover the sparse vector.

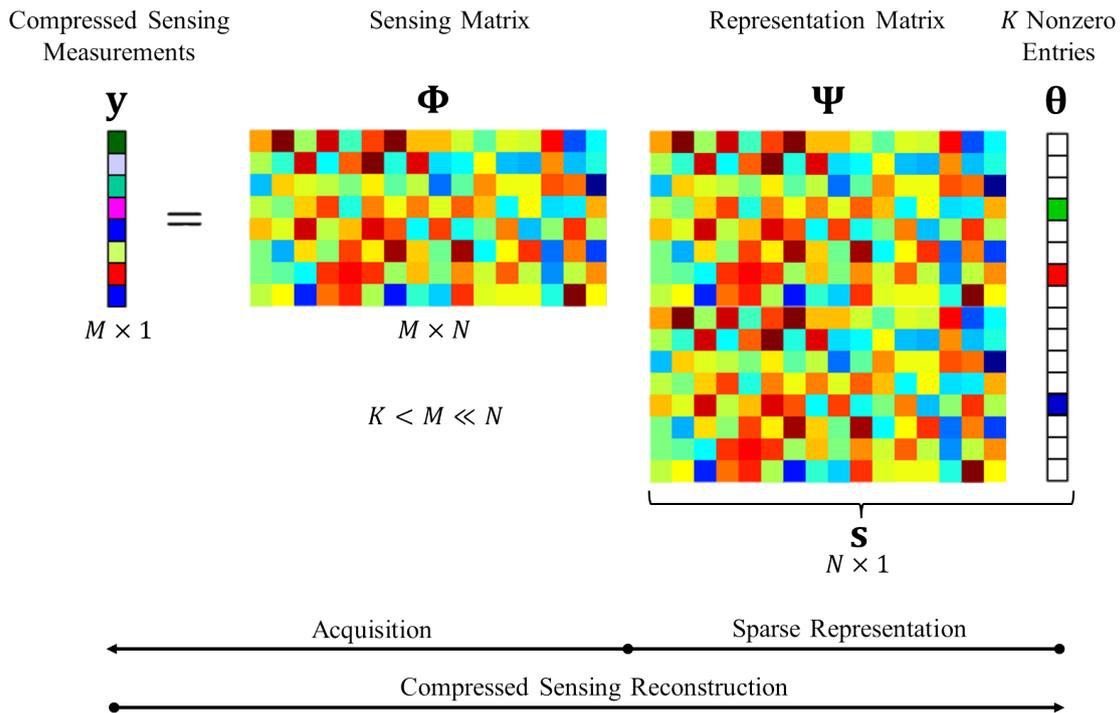


FIGURE 2.2: Compressed sensing scheme. The speech signal \mathbf{s} of length N is sparsified by projecting it on the representation basis Ψ that can be Fourier basis, wavelets, ... As a result, we obtain the sparse vector θ (of length N) that contains only K non null values, so we say that it is K -sparse. The acquisition is performed by multiplying the measurement matrix Φ of length $M \times N$ by the signal \mathbf{s} . Hence, we obtain the acquired signal \mathbf{y} (measurement vector) that contains only M measurements, such that $K < M \ll N$

In this section we describe the basic concepts related to CS.

2.3.1 Sparsity

Compressed sensing is based on the fact that most real-life signals such as images, speech, music, radar signals, and ultrasound signals are well approximated as a linear combination of only a small number of relevant coefficients from a suitable basis or dictionary. This signal property is called *sparsity*, and it is prevalent for sparse recovery. Mathematically, this is to express a signal \mathbf{s} with a sparse coefficients vector θ containing only a few nonzero values (K entries).

Let us consider a real¹ signal $\mathbf{s} \in \mathbb{R}^N$ - speech in our case². Using an appropriate *sparsity basis* (also called *representation basis*) $\Psi = \{\psi_1, \psi_2, \psi_3, \dots, \psi_N\}$, the speech

¹Although our discussion is restricted to real vectors, all ideas discussed here extend trivially also to vectors with entries in \mathbb{C}

²As our applications use speech signals, the latter is considered in the whole document.

signal \mathbf{s} can be represented by a linear combination as:

$$\mathbf{s} = \sum_{i=1}^N \theta_i \boldsymbol{\psi}_i \quad (2.1)$$

where $\boldsymbol{\theta} = [\theta_1, \theta_2, \theta_3, \dots, \theta_N]^T$ is the sparse coefficients vector of the speech signal \mathbf{s} of length N in the basis $\boldsymbol{\Psi}$, $\theta_i = \langle \mathbf{s}, \boldsymbol{\psi}_i \rangle$; $\langle \cdot, \cdot \rangle$ denotes the scalar product. \mathbf{s} is more simply written as

$$\mathbf{s} = \boldsymbol{\Psi} \boldsymbol{\theta} \quad (2.2)$$

where $\boldsymbol{\Psi}$ is a $N \times N$ matrix with $\boldsymbol{\psi}_i$ as column i .

Strict Sparsity and Compressibility

If the vector $\boldsymbol{\theta}$ contains at most $K \ll N$ nonzero entries, then $\boldsymbol{\theta}$ is called *K-sparse*.

Definition 1. The *support* of a vector $\boldsymbol{\theta}$ is the set of positions of its nonzero entries

$$\text{supp}(\boldsymbol{\theta}) = \{i \in \{1, 2, \dots, N\} : \theta_i \neq 0\} \quad (2.3)$$

Definition 2. (Haltmeier et al., 2016) The *strict K-sparsity* is achieved when

$$\text{card}(\text{supp}(\boldsymbol{\theta})) \leq K \ll N \quad (2.4)$$

where $\text{card}(\cdot)$ denotes the cardinality of the set. The ℓ_0 norm³ can be used to count the total number of nonzero elements in a vector, $\text{card}(\text{supp}(\boldsymbol{\theta})) = \|\boldsymbol{\theta}\|_0$.

It is more appropriate to write $\|\boldsymbol{\theta}\|_0$ as $\|\boldsymbol{\theta}\|_0^0$ because (Foucart and Rauhut, 2013)

$$\|\boldsymbol{\theta}\|_p^p = \sum_{i=1}^N |\theta_i|^p \xrightarrow{p \rightarrow 0} \sum_{i=1}^N \mathbb{1}_{\{\theta_i \neq 0\}} = \text{card}(\text{supp}(\boldsymbol{\theta})) \quad (2.5)$$

considering $\mathbb{1}_{\{\theta_i \neq 0\}} = 1$ if $\theta_i \neq 0$ and $\mathbb{1}_{\{\theta_i \neq 0\}} = 0$ if $\theta_i = 0$. The ℓ_0 norm satisfies $\|\boldsymbol{\theta}\|_0 = \lim_{p \rightarrow 0} \|\boldsymbol{\theta}\|_p^p$, where

$$\|\boldsymbol{\theta}\|_p = \sqrt[p]{\sum_{i=1}^N |\theta_i|^p}, \quad p > 0 \quad (2.6)$$

stands for the ℓ_p norm. Indeed, if $p \geq 1$, $\|\cdot\|_p$ is a norm, and if $0 < p < 1$, it is a quasi-norm.

³Mathematically, $\|\cdot\|_0$ is not a norm, but it is commonly (and abusively) called ℓ_0 norm. Hence, we keep this appellation in the rest of this thesis.

Real-life signals are often not K -sparse in the strict sense, for example, when signals are affected by noise, coefficients with very large and very small values exist and make the value of K large. Hence, we prefer a weaker concept, which is *compressibility*⁴. In this case, to achieve a good sparse approximation, one can eliminate small coefficients without much loss (Candès and Wakin, 2008) using, for example, thresholding to obtain a K -sparse representation. Let us define \mathbf{s}_K by keeping only the K largest values of $\boldsymbol{\theta}$. Consequently, $\mathbf{s}_K = \mathbf{\Psi}\boldsymbol{\theta}_K$ where $\boldsymbol{\theta}_K$ is the vector of coefficients θ_i with all $(N - K)$ smallest values set to zero. $\boldsymbol{\theta}_K$ is called *the best K -term approximation*. It is the nearest K -sparse vector to $\boldsymbol{\theta}$ yielding the minimum error of best K -term approximation defined as follows.

Definition 3. (Foucart and Rauhut, 2013) For $p > 0$, the ℓ_p error of best K -term approximation to a vector $\boldsymbol{\theta}$ is given by

$$\sigma_K(\boldsymbol{\theta})_p = \inf\{\|\boldsymbol{\theta} - \boldsymbol{\theta}_K\|_p, \quad \boldsymbol{\theta}_K \text{ is } K\text{-sparse}\} \quad (2.7)$$

Note that when $\|\boldsymbol{\theta}\|_0 \leq K$ then $\boldsymbol{\theta} = \boldsymbol{\theta}_K$ (strict sparsity), and when the distance between $\boldsymbol{\theta}$ and $\boldsymbol{\theta}_K$ is small then $\boldsymbol{\theta}$ is *compressible* (Amin, 2015). Vectors whose sorted entries decay quickly are considered to be *compressible* (Amin, 2015). Hence, we define *compressibility* as follows.

Definition 4. (Frigo, 2014) A vector $\boldsymbol{\theta}$ is said to be *compressible* if there exist two constants $c > 0$ and $r > 0$ that verify the following inequality

$$\|\boldsymbol{\theta} - \boldsymbol{\theta}_K\|_p \leq cK^{-r} \quad (2.8)$$

Besides the ℓ_0 norm, other measures can be used to calculate a number that describes the sparsity of a vector. Commonly used sparsity measures are described in the subsequent section.

Sparsity Measures

Sparsity is commonly defined in view of the ℓ_0 norm as the number of nonzero entries. Nevertheless, in many real situations (e.g., in noisy case), this definition may not be practical. An alternative definition says that a signal is considered sparse if its energy is concentrated in a small number of coefficients (Zonoobi, Kassim, and Venkatesh, 2011). Based on this interpretation, various sparsity measures were proposed in the literature (Hurley and Rickard, 2009; Karvanen and Cichocki, 2003;

⁴In this thesis, the term *sparse* signal is used interchangeably with *compressible*.

Rickard and Fallon, 2004). Several commonly used measures are presented in this section, which are: ℓ_0 , ℓ_0^ϵ , ℓ_1 , ℓ_p ($0 < p < 1$), ℓ_p ($-1 < p < 0$), $\frac{\ell_2}{\ell_1}$, Hoyer measure, $\tanh_{a,b}$, \log , κ_4 , $\kappa_{a,b}$, u_δ^0 , and Gini index.

The most commonly used and studied sparsity measures are those based on the ℓ_p norm.

The ℓ_0 measure is a traditional sparsity measure. Its drawbacks are that its derivative does not contain information, and its poor behaviour in many practical situations such in the noisy case (Hurley and Rickard, 2009).

In the latter example case, ℓ_0^ϵ can be used. This measure counts the number of coefficients greater than a threshold ϵ . It is defined, for a sparse vector θ , as

$$\|\theta\|_{0,\epsilon} = \text{card}(\{i \in \{1, 2, \dots, N\} : |\theta_i| \geq \epsilon\}) \quad (2.9)$$

The parameter ϵ depends on the noise variance, which is unknown. Hence determining its value is an open problem. Another practical problem with this measure is that the optimization with gradient methods using this norm is not easy as it is non-differentiable (Karvanen and Cichocki, 2003).

ℓ_p with $0 < p < 1$ is often used instead of ℓ_0^ϵ (Xu et al., 2007). In order to imitate the ℓ_0 norm, small values of p , e.g. 0.1 or 0.01, should be used (Karvanen and Cichocki, 2003).

One of the most widely used sparsity measures is the ℓ_1 norm (ℓ_p with $p = 1$). The latter considers that large coefficients are more important than small coefficients. It approximates the ℓ_0 measure, and under some settings, it can be used to find the support of the ℓ_0 solution. This measure is easily calculated by linear programming; hence its usage in many optimization problems (Candès and Tao, 2005; Donoho and Tsai, 2008).

ℓ_p with $-1 < p < 0$ are also used for measuring sparsity (Rao and Kreutz-Delgado, 1999).

The $\frac{\ell_2}{\ell_1}$ measure or its normalized version (Hoyer measure) can be used as sparsity measure. The Hoyer measure is defined as (Hoyer, 2004)

$$K = \frac{\sqrt{N} - \frac{\ell_1}{\ell_2}}{\sqrt{N} - 1} \quad (2.10)$$

Instead of ℓ_p with $0 < p < 1$, $\tanh_{a,b}$ is sometimes utilized (Hurley and Rickard, 2009). This measure says that a representation with one large component is sparser

than another with two smaller ones. $\tanh_{a,b}$ is defined as follows (Karvanen and Cichocki, 2003; Rickard and Fallon, 2004)

$$\tanh_{a,b}(\boldsymbol{\theta}) = \sum_{i=1}^N \tanh(|a\boldsymbol{\theta}_i|^b) \quad (2.11)$$

where a and b are positive constants. The difference between $\tanh_{a,b}$ and the ℓ_p norm is that the former is limited to the range $[0, 1]$ (Karvanen and Cichocki, 2003). In order to imitate the ℓ_0^ϵ measure, we must choose $b > 1$, e.g. $b = 2$ or $b = 3$.

It is also possible to measure sparsity using the log measure, which enforces sparsity outside some range (Rickard and Fallon, 2004). It is defined as follows (Rickard and Fallon, 2004)

$$\log(1 + \boldsymbol{\theta}^2) = \sum_{i=1}^N (1 + \boldsymbol{\theta}_i^2) \quad (2.12)$$

The kurtosis κ_4 , which measures the peakedness of a distribution, can also be used as sparsity measure (Karvanen and Cichocki, 2003). It is defined as

$$\kappa_4(\boldsymbol{\theta}) = \frac{\sum_{i=1}^N \boldsymbol{\theta}_i^4}{(\sum_{i=1}^N \boldsymbol{\theta}_i^2)^2} \quad (2.13)$$

One can use the generalized kurtosis (or Gray's variable norm) $\kappa_{a,b}$ instead of the kurtosis. It is defined as (Karvanen and Cichocki, 2003)

$$\kappa_{a,b}(\boldsymbol{\theta}) = \frac{\sum_{i=1}^N |\boldsymbol{\theta}_i|^a}{(\sum_{i=1}^N |\boldsymbol{\theta}_i|^{a/b})^b} \quad (2.14)$$

a and b are positive, e.g. ($b = 2, a = 1, 3, 4, 6$). The kurtosis is a special case obtained when $a = 4$ and $b = 2$ (Karvanen and Cichocki, 2003).

Another sparsity measure is u_δ^0 (Hurley and Rickard, 2009; Rickard and Fallon, 2004; Karvanen and Cichocki, 2003). It is defined for the ordered data $\boldsymbol{\theta}_{(1)} \leq \boldsymbol{\theta}_{(2)} \leq \dots \leq \boldsymbol{\theta}_{(N)}$ as (Rickard and Fallon, 2004; Karvanen and Cichocki, 2003)

$$u_\delta^0(\boldsymbol{\theta}) = \min_{i,j} (\boldsymbol{\theta}_{(j)} - \boldsymbol{\theta}_{(i)}) \text{ subject to } \frac{j-i}{N} \geq \delta \text{ and } \boldsymbol{\theta}_{(i)} \leq 0 \leq \boldsymbol{\theta}_{(j)} \quad (2.15)$$

where $(1), (2), \dots, (N)$ are the new indices after sorting the coefficients. u_δ^0 measures the smallest range around the origin which contains a certain percentage of the data.

This is achieved by sorting the data and determining the minimum difference between the largest and smallest sample in a range containing the specified percentage (δ) of data points (Hurley and Rickard, 2009). Depending on the application, the interesting value of δ might be for example 0.5, 0.75, or 0.99. $\delta = 0.99$ means that 99% of values are required to be concentrated on a small area (Karvanen and Cichocki, 2003).

Another sparsity measure is called Gini index (GI) (Hurley and Rickard, 2009; Zonoobi, Kassim, and Venkatesh, 2011; Rickard and Fallon, 2004). It is computed, after sorting the coefficients in ascended order, as

$$GI(\boldsymbol{\theta}) = 1 - 2 \sum_{i=1}^N \frac{|\boldsymbol{\theta}_{(i)}|}{\|\boldsymbol{\theta}\|} \left(\frac{N - i + \frac{1}{2}}{N} \right) \quad (2.16)$$

where $|\boldsymbol{\theta}_{(1)}| \leq |\boldsymbol{\theta}_{(2)}| \leq \dots \leq |\boldsymbol{\theta}_{(N)}|$. (1), (2), ..., (N) are the new indices after the sorting operation.

Other sparsity measures exist; namely, pq -mean, normalized kurtosis, Shannon entropy measure and modified versions of it (Hurley and Rickard, 2009).

How to Sparsify Signals?

Sparsity is the first fundamental constraint for recovery in CS. Most signals do not have a sparse representation in the time domain, but they are sparse when expressed in a suitable sparsifying basis Ψ or dictionary \mathbf{D} .

Sparsifying Transform Speech signals vary highly in time, so they are naturally not sparse in the time domain. Several transforms can be used to sparsify the speech signals such as discrete cosine transform, discrete Fourier transform, quadratic time-frequency distributions, wavelets, *etc.* Other nonorthogonal transforms include total variation and curvelet transform. As mentioned before, for compressible signals, small coefficients can be eliminated (using thresholding) to reinforce sparsity.

These transforms are analytic. They have been widely investigated and examined in applications for their easy and fast implementations, compared to the learned dictionaries, described next.

Dictionary Learning Dictionary learning aims at finding a sparse representation of the input data in the form of a linear combination of basic elements called atoms. These atoms compose the dictionary. Useful dictionaries can be generated from various training sources using modern training approaches such as the method of optimal

directions (Engan, Aase, and Husøy, 2000), K-singular value decomposition (K-SVD) (Xu et al., 2017), and many others. Dictionaries can be formed from scratch, with no structural restrictions whatsoever, or from certain distributions with parameters learned from the training set. The training and usage of these dictionaries, especially those without any structures, are computationally expensive, but these dictionaries tend to do better on complex signals and complicated tasks for which no analytic dictionaries have been designed (Han, Li, and Yin, 2013). K-SVD is a widely used dictionary learning algorithm. It is a generalization of k-means clustering technique. K-SVD is an iterative method alternating between two steps: (a) sparse coding, which uses orthogonal matching pursuit (OMP) algorithm to find the sparse representation of the input signal, and (2) dictionary updating using singular value decomposition.

2.3.2 Acquisition

The compressed sensing acquisition of the speech signal \mathbf{s} corresponds to collecting a small number of linear measurements of \mathbf{s} . This can be represented as:

$$\mathbf{y}_i = \langle \mathbf{s}, \boldsymbol{\phi}_i \rangle, \quad i = 1, \dots, M \quad (2.17)$$

where $\boldsymbol{\phi}_i$ is the i^{th} row of the *sensing matrix* $\boldsymbol{\Phi}$ of dimensions $M \times N$ with $K < M \ll N$, and M is the number of measurements required for reconstruction. $\langle \cdot, \cdot \rangle$ denotes the inner product. The signal \mathbf{y} , called the *observation vector* (or the *measurement vector*) can also be written (in a matrix form) as:

$$\mathbf{y} = \boldsymbol{\Phi} \mathbf{s} = \boldsymbol{\Phi} \boldsymbol{\Psi} \boldsymbol{\theta} = \mathbf{A} \boldsymbol{\theta} \quad (2.18)$$

where $\boldsymbol{\theta}$ is the vector of sparse coefficients of the signal. \mathbf{y} is said to contain *compressive measurements* of \mathbf{s} .

Perfect reconstruction requires that $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$ must obey two conditions known as (a) incoherence condition and (b) restricted isometry property (Candès, Romberg, and Tao, 2006).

Incoherence

Let $(\boldsymbol{\Phi}, \boldsymbol{\Psi})$ be a pair of orthobases of \mathbb{R}^N . The first basis $\boldsymbol{\Phi}$ (also known as *sensing matrix*) is used for sensing or measuring the speech signal \mathbf{s} (see Eq. (2.17)), and the second basis $\boldsymbol{\Psi}$ (also known as the *sparsity* or *representation matrix*) is used to sparsely represent \mathbf{s} . It may not be sparse in the time domain, but its expansion in the basis $\boldsymbol{\Psi}$ may give Eq. (2.1).

The incoherence means that the mutual coherence between the sensing matrix Φ and the sparse basis Ψ is small. The mutual coherence measures the largest correlation between any two elements of Φ and Ψ . In other words, it is the maximum absolute value of the cross-correlations between the columns of Φ and Ψ . It can be defined as (Candès and Tao, 2006)

$$\mu(\Phi, \Psi) = \alpha \max_{1 \leq k, j \leq N} | \langle \phi_k, \psi_j \rangle | \quad (2.19)$$

where ϕ_k and ψ_j are the k th and j th columns of Φ and Ψ , respectively. According to various authors, α can take the values N , \sqrt{N} , or a value normalized to one. The value of α depends of the algebraic properties of the chosen matrices Φ and Ψ . In this thesis the selected value is: $\alpha = \sqrt{N}$, which gives a range of variation of the coherence between 1 and \sqrt{N} (Donoho and Huo, 2001; Candès and Wakin, 2008) For incoherent matrices, μ is close to the lower bound.

The coherence serves as a rough characterization of the degree of similarity (or “correlation”) between the sparsity and measurement systems. The value $\mu = 1$ is the minimum (lowest value) of the coherence, which means that the rows of the measurement matrix Φ must be spread out in the Ψ domain. In other words, the lower the coherence, the greater the incoherence. This idea appears in CS when choosing the sensing matrix: the sensing matrix must be incoherent with the basis in which sparsity is expressed. For more clarification, let us consider two particular examples of matrices Φ and Ψ :

- Let Φ be the sensing matrix corresponding to classical sampling in time or space, and Ψ be the Fourier basis. The time-frequency pair satisfies $\mu(\Phi, \Psi) = 1$, and therefore, there is maximum incoherence. The notion of incoherence extends the duality between time and frequency: a very localized signal in time is spread in frequency (principle of uncertainty).
- The sensing matrix Φ is chosen as a Gaussian random matrix whose elements are mutually independent and the representation matrix Ψ is an orthogonal wavelet basis (e.g. Daubechies wavelet of order one or four). This choice is motivated by the fact that if the measurement matrix is random, then it is largely incoherent with any fixed basis Ψ (Candès and Tao, 2006). For the latter example μ is close to $\sqrt{2 \log N}$ (Candès and Wakin, 2008).

Restricted Isometry Property

This is a key issue in CS when studying its general robustness: the so called restricted isometry property (RIP). It requires that all subsets of K columns taken from the

product matrix \mathbf{A} (defined as $\mathbf{A} = \mathbf{\Phi\Psi}$) are in fact nearly orthogonal (they can't be exactly orthogonal since we have more columns than rows).

It is expressed as:

$$1 - \delta_K \leq \frac{\|\mathbf{\Phi\Psi\theta}\|_2^2}{\|\theta\|_2^2} \leq 1 + \delta_K \quad (2.20)$$

where δ_K is the RIP constant of order K , and $\|\cdot\|_2$ denotes the ℓ_2 norm. The matrix \mathbf{A} is said to satisfy the K -restricted isometry property with restricted isometry constant δ_K .

The RIP is the sufficient condition for CS matrices that guarantees the performance of the signal reconstruction in terms of efficiency and robustness (Candès and Wakin, 2008). As random matrices with independent identically distributed entries fulfill the RIP condition with high probability, they are the most used measurement matrices in CS (Baraniuk et al., 2008).

Since it has been demonstrated (Candès and Tao, 2005; Baraniuk et al., 2008; Candès and Romberg, 2007; Candès and Tao, 2006) that the Gaussian random matrix fills the two conditions when used with any other basis, it is considered as a universal sensing matrix that can be used with any sparse basis (including those trained using dictionary learning); hence its wide utilization. Gaussian sensing matrix is used in this work.

2.3.3 Signal Reconstruction

The recovery aims to find the sparsest θ subject to Eq. (2.18).

It is possible under the two aforementioned conditions to reconstruct θ (of length N) from $M \geq O(K \log(N/K))$ measurements (Candès and Wakin, 2008) (i.e., the observation vector $\mathbf{y} = \mathbf{A}\theta$ of length M). The matrix $\mathbf{A} = \mathbf{\Phi\Psi}$ is of dimension $M \times N$ with $M < N$. Therefore, \mathbf{A} has a rank lower than N and is not reversible. This implies that there is an infinity of solutions $\hat{\theta}$ such as $\mathbf{A}\hat{\theta} = \mathbf{A}\theta$ with $\hat{\theta} \neq \theta$. Solving the problem is to determine which of these solutions corresponds to θ .

The Moore-Penrose pseudo-inverse is a conventional approach to obtain the minimum ℓ_2 norm solution. It is defined as

$$\begin{aligned} (P_2) : \min_{\theta} \|\theta\|_2 \quad \text{subject to } \mathbf{A}\theta = \mathbf{y} \\ \Leftrightarrow \min_{\theta} \left(\sum_{i=1}^N |\theta_i|^2 \right)^{1/2} \quad \text{subject to } \mathbf{A}\theta = \mathbf{y} \end{aligned} \quad (2.21)$$

If the matrix \mathbf{A} is of rank M , the solution is given by the equation below

$$\hat{\boldsymbol{\theta}} = \mathbf{A}^+ \mathbf{y} = \mathbf{A}^T (\mathbf{A} \mathbf{A}^T)^{-1} \mathbf{y} \quad (2.22)$$

where \mathbf{A}^+ is the Moore-Penrose pseudo-inverse, and \mathbf{A}^T is the transpose of \mathbf{A} .

Indeed, minimizing the ℓ_2 norm does not promote the sparsity of the solution since such norm does not disadvantage small coefficients. So, the hypothesis of the existence of a sparse decomposition of the signal is not considered. To exploit this prior information (the sparsity), one must use an approach that disadvantages non sparse solutions. The natural approach is, therefore, to minimize the ℓ_0 pseudo norm of the solution, which results in the following problem

$$\begin{aligned} (P_0) : \min_{\boldsymbol{\theta}} \|\boldsymbol{\theta}\|_0 \quad \text{subject to } \mathbf{A}\boldsymbol{\theta} = \mathbf{y} \\ \Leftrightarrow \min_{\boldsymbol{\theta}} \left(\sum_{i=1}^N |\theta_i|^0 \right) \quad \text{subject to } \mathbf{A}\boldsymbol{\theta} = \mathbf{y} \end{aligned} \quad (2.23)$$

where $(\sum_{i=1}^N |\theta_i|^0)$ is the cardinal of $\boldsymbol{\theta}$, i.e., the total number of non-zero elements in a vector (see section 2.3.1). However, looking for such solution is an NP-hard problem. To overcome the feasibility problem of finding a solution as defined in (P_0) , ℓ_0 constraint is replaced by ℓ_1 norm; in fact, ℓ_1 norm can be seen as a relaxation of ℓ_0 pseudo norm, so it ensures sparsity. In some cases, solving (P_1) gives the exact solution of (P_0) (Candès and Wakin, 2008):

$$\begin{aligned} (P_1) : \min_{\boldsymbol{\theta}} \|\boldsymbol{\theta}\|_1 \quad \text{subject to } \mathbf{A}\boldsymbol{\theta} = \mathbf{y} \\ \Leftrightarrow \min_{\boldsymbol{\theta}} \left(\sum_{i=1}^N |\theta_i|^1 \right) \quad \text{subject to } \mathbf{A}\boldsymbol{\theta} = \mathbf{y} \end{aligned} \quad (2.24)$$

The advantage of the formulation (P_1) is that it can be interpreted as a convex optimization problem, which can be solved by linear programming if the coefficients are real (Candès and Wakin, 2008).

After estimating the sparse coefficient vector, one can synthesise the estimated speech as $\hat{\mathbf{s}} = \boldsymbol{\Psi} \hat{\boldsymbol{\theta}}$. For orthonormal bases $\boldsymbol{\Psi}$, the error in the signal domain is equal to the error in the coefficient domain (Amin, 2015, Chapter 1):

$$\|\mathbf{s} - \hat{\mathbf{s}}\|_2 = \|\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\|_2 \quad (2.25)$$

The compressive sensing reconstruction algorithms are classified under six approaches summarized in the following subsections.

Convex Optimization Approach

If a signal \mathbf{s} is sufficiently sparse (in the basis $\mathbf{\Psi}$), it can be perfectly reconstructed by minimizing the ℓ_1 norm of the solution. This approach is known as “basis pursuit” (Candès and Wakin, 2008). The purpose of this technique is to find vectors with the smallest norm ℓ_1 (Eq. (2.24)). If the measurements are affected by an additive noise, i.e.,

$$\mathbf{y} = \mathbf{\Phi}\mathbf{s} + \boldsymbol{\omega} \quad (2.26)$$

where $\boldsymbol{\omega} \in \mathbb{R}^M$ is an unknown noise such that $\|\boldsymbol{\omega}\|_2 \leq \varepsilon$, the minimization problem (P_1) needs to be changed to the following problem known as basis pursuit denoising (BPDN) (Yang et al., 2013; Amin, 2015)

$$\text{(BPDN): } \min_{\boldsymbol{\theta}} \|\boldsymbol{\theta}\|_1 \quad \text{subject to } \|\mathbf{A}\boldsymbol{\theta} - \mathbf{y}\|_2^2 \leq \varepsilon \quad (2.27)$$

where ε is a user-selected parameter related to noise power. For $\varepsilon > 0$ carefully chosen and for a particular parameter $\lambda > 0$, this problem can be solved by the unconstrained optimization problem defined as

$$\min \frac{1}{2} \|\mathbf{A}\boldsymbol{\theta} - \mathbf{y}\|_2^2 + \lambda \|\boldsymbol{\theta}\|_1 \quad (2.28)$$

which is a Lagrangian form of the BPDN problem. The parameter λ assures a trade off between the sparsity of the solution and the distance between $\mathbf{A}\hat{\boldsymbol{\theta}}$ and $\mathbf{y} = \mathbf{A}\boldsymbol{\theta}$. More sparse solution $\hat{\boldsymbol{\theta}}$ is obtained using large value of λ , and small value of $\|\mathbf{y} - \mathbf{A}\hat{\boldsymbol{\theta}}\|_2$ is achieved using small value of λ .

In addition to BPDN, other optimization based strategies have been proposed for robust sparse recovery from noisy measurements, namely, Dantzig selector and total variation denoising.

As stated before, basis pursuit can be cast as linear program if \mathbf{s} is real. When \mathbf{s} is complex valued, basis pursuit can be cast as a second-order cone program (SOCP). Popular solvers for convex optimization problems are (Amin, 2015) simplex algorithm, interior-point algorithm, fixed point continuation (FCP), gradient projection for sparse representation (GPSR), Bregman iteration algorithm, *etc.*

To solve the basis pursuit problem, popular software packages have been designed. Major work along these lines has included ℓ_1 -MAGIC (Candès and Romberg, 2015), CVX (Grant and Boyd, 2018; Grant and Boyd, 2008), homotopy algorithm (Efron et al., 2004; Osborne, Presnell, and Turlach, 2000), linearized Bregman iterations (Yin et al., 2008), spectral projected-gradient algorithm (SPGL1) (Berg and

Friedlander, 2018; Berg and Friedlander, 2008), and an algorithm based on the alternating direction method (YALL1) (Zhang, Yang, and Yin, 2011; Yang and Zhang, 2011). To solve the BPDN problem, the standard methods for convex optimization can be used, but better performance can be achieved using SPGL1, an algorithm inspired by Nesterov (NESTA) (Becker, Bobin, and Candès, 2010; Becker, Bobin, and Candès, 2011), a formulation of the problem in conic form (TFOCS) (Becker, Candès, and Grant, 2013; Becker, Candès, and Grant, 2011), the sparse modeling software (SPAMS) toolbox (Mairal, 2017; Mairal et al., 2009; Mairal et al., 2010), and YALL1 (Zhang, Yang, and Yin, 2011; Yang and Zhang, 2011). To solve the unconstrained optimization problem presented by Eq. (2.28), large number of algorithms can be used, including the least angle regression (LARS) homotopy algorithm (Efron et al., 2004; Malioutov, Cetin, and Willsky, 2005; Osborne, Presnell, and Turlach, 2000), a gradient projection technique (GPSR) (Figueiredo, Nowak, and Wright, 2009; Figueiredo, Nowak, and Wright, 2007), a fixed-point continuation method (FPC) (Hale, Yin, and Zhang, 2008), an approach utilizing linearized Bregman iterations (Yin et al., 2008), a coordinate descent technique (Friedman, Hastie, and Tibshirani, 2010), the SPAMS toolbox (Mairal, 2017; Mairal et al., 2009; Mairal et al., 2010), and YALL1 (Zhang, Yang, and Yin, 2011; Yang and Zhang, 2011). Some of these techniques are limited to real vectors and matrices (e.g., ℓ_1 -MAGIC, homotopy algorithm, and linearized Bregman iterations); others, such as CVX and YALL1, can accommodate complex matrices and vectors.

Greedy Approach

Differently from the global optimization method (i.e., convex optimization approach), the greedy approach is a step-by-step iterative method. The algorithms of this approach iteratively approximate the coefficients and the support of the original signal, by choosing only the columns (called atoms) of the reconstruction matrix ($\mathbf{A} = \Phi\Psi$) that have a high correlation with the measurements. The selected atoms are, generally, not included in subsequent iterations; thus the computational complexity is reduced. The solution is approached in a greedy manner; hence the name. Their advantages are to be very fast, easy to implement, and with low computational complexity. The most famous greedy algorithm is orthogonal matching pursuit (OMP) (Pati, Rezaiifar, and Krishnaprasad, 1993; Tropp and Gilbert, 2007) which is an improved successor of the matching pursuit. Other remarkable algorithms are regularized OMP (Needell and Vershynin, 2009), compressive sampling matching pursuit (CoSaMP) (Needell and Tropp, 2009), and greedy stagewise OMP (Donoho et al., 2012). Greedy algorithms are categorized into two classes: serial (e.g., OMP) and parallel (e.g., CoSaMP).

Serial Greedy Algorithms Algorithms of this category select only one atom in each iteration, and compute the corresponding non-zero entry of the solution (Fig. (2.3)). Popular examples of these algorithms are matching pursuit (Mallat and Zhang, 1993), OMP (Pati, Rezaiifar, and Krishnaprasad, 1993; Tropp and Gilbert, 2007), and gradient pursuit (Blumensath and Davies, 2008). The most famous greedy algorithm is OMP. Its basic steps are summarized as follows (see also Fig. (2.3) and Algorithm 1)

- Initialize the index set Λ to an empty set, the sparse vector $\hat{\boldsymbol{\theta}}$ and the iteration number i to 0, and the residual vector \mathbf{r} to \mathbf{y} .
- Search atom in the reconstruction matrix \mathbf{A} that is maximally correlated with \mathbf{r} , and update the index set Λ by including the position of this atom.
- Update the solution vector by using the least square method to find the sparse vector, with support Λ , that best fits the measurements.
- Update the residual by subtracting $\mathbf{A}\boldsymbol{\theta}^{(i+1)}$ from the measurement vector \mathbf{y} .

Algorithm 1 Orthogonal matching pursuit (OMP)

Input: \mathbf{A} the reconstruction matrix, \mathbf{y} the measurements, and the stopping criterion

Output: The sparse vector $\hat{\boldsymbol{\theta}}$

1: **Initialize:**

$$\Lambda^{(0)} \leftarrow \emptyset, \boldsymbol{\theta}^{(0)} \leftarrow 0, \mathbf{r}^{(0)} \leftarrow \mathbf{y}, i \leftarrow 0$$

2: **while** stopping criterion not met **do**

3: $\mathbf{g}^{(i)} \leftarrow \mathbf{A}^T \mathbf{r}^{(i)}$

4: $j \leftarrow \arg \max_j |\mathbf{g}^{(i)}|$ ▷ choose one value for j if multiple exist

5: $\Lambda^{(i+1)} \leftarrow \Lambda^{(i)} \cup j$

6: $\boldsymbol{\theta}^{(i+1)} \leftarrow \arg \min_{\mathbf{z}} \|\mathbf{y} - \mathbf{A}\mathbf{z}\|_2, \text{supp}(\mathbf{z}) \subseteq \Lambda^{(i+1)}$

7: $\mathbf{r}^{(i+1)} \leftarrow \mathbf{y} - \mathbf{A}\boldsymbol{\theta}^{(i+1)}$

8: $i \leftarrow i + 1$

9: **end while**

10: **return** $\hat{\boldsymbol{\theta}}$

The stopping criterion can be based on the sparsity K or the residual. Note that these steps are the same for other serial greedy algorithms except the solution updating.

A weakness of OMP is that if an incorrect index has been selected in an iteration, it remains as it is in the following iterations. Hence, K iterations are not enough. A

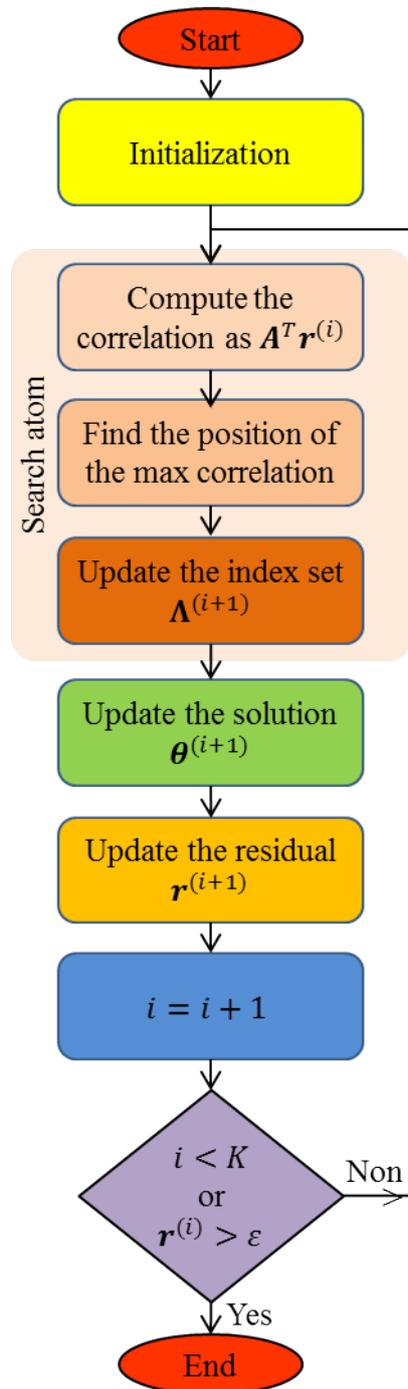


FIGURE 2.3: Serial greedy algorithms steps

possible solution is to increase the number of iterations or to use the compressive sampling matching pursuit algorithm which is classified in the following category.

Parallel Greedy Algorithms Algorithms of this category select K or multiple of K atoms at a time instead of only one atom; hence the name of parallel greedy algorithms. The other steps are the same as described for serial algorithms (Fig. (2.3)). Examples of these algorithms are CoSaMP (Needell and Tropp, 2009) and subspace pursuit (Dai and Milenkovic, 2009). Their basic steps are summarized as follows (Algorithm 2)

- Initialize the sparse vector $\hat{\boldsymbol{\theta}}$ and the iteration number i to 0, and the residual vector \mathbf{r} to \mathbf{y} .
- Search $2K$ atoms in the reconstruction matrix \mathbf{A} that are maximally correlated with \mathbf{r} , and update the index set Λ by adding the position of these atoms to the K previously selected.
- Update the solution vector by (a) using the least square method to find the sparse vector, with support Λ , that best fits the measurements; then, (b) keeping the largest K values (using hard thresholding operator).
- Update the residual by subtracting $\mathbf{A}\boldsymbol{\theta}^{(i+1)}$ from the measurement vector \mathbf{y} .

Algorithm 2 Compressive sampling matching pursuit (CoSaMP)

Input: \mathbf{A} the reconstruction matrix, \mathbf{y} the measurements, K the sparsity level, and the stopping criterion

Output: The sparse vector $\hat{\boldsymbol{\theta}}$

1: **Initialize:**

$$\Lambda^{(0)} \leftarrow \emptyset, \boldsymbol{\theta}^{(0)} \leftarrow \mathbf{0}, \mathbf{r}^{(0)} \leftarrow \mathbf{y}, i \leftarrow 0$$

2: **while** stopping criterion not met **do**

3: $\mathbf{g}^{(i)} \leftarrow \mathbf{A}^T \mathbf{r}^{(i)}$

4: Identify $\mathbf{g}_{2K}^{(i)}$ (by keeping only the largest $2K$ entries of $\mathbf{g}^{(i)}$)

5: $\Lambda^{(i+1)} \leftarrow \text{supp}(\boldsymbol{\theta}^{(i)}) \cup \text{supp}(\mathbf{g}_{2K}^{(i)})$

6: $\mathbf{z} \leftarrow \arg \min_{\mathbf{z}} \|\mathbf{y} - \mathbf{A}\mathbf{z}\|_2, \text{supp}(\mathbf{z}) \subseteq \Lambda^{(i+1)}$

7: $\boldsymbol{\theta}^{(i+1)} \leftarrow \mathbf{z}_K$ ▷ keep the largest K elements of \mathbf{z}

8: $\mathbf{r}^{(i+1)} \leftarrow \mathbf{y} - \mathbf{A}\boldsymbol{\theta}^{(i+1)}$

9: $i \leftarrow i + 1$

10: **end while**

11: **return** $\hat{\boldsymbol{\theta}}$

Indeed, CoSaMP overestimates the size of the support set, and then estimates the signal coefficients and uses these estimates to prune back the support. These steps are repeated until small residual is achieved. As stated before, unlike serial algorithms these algorithms can remove the wrong atoms selected during previous iterations. Therefore, they are more powerful.

Thresholding Approach

This approach updates the solution $\hat{\theta}$ greedily using some thresholding operation. It operates on K atoms of \mathbf{A} simultaneously. Popular examples of algorithms under this category are iterative hard thresholding (IHT) (Blumensath and Davies, 2009), iterative soft thresholding (Daubechies, Defrise, and Mol, 2004), approximate message passing (Donoho, Maleki, and Montanari, 2009), *etc.* Note that CoSaMP and subspace pursuit are classified also in this category as they use a hard thresholding operator.

Iterative hard thresholding (Blumensath and Davies, 2009) refine iteratively the sparse vector estimate. It uses a non-linear thresholding $Hthr_K(\cdot)$ operator that keep the K largest coefficients and set the others to zero. The steps of IHT are described as follows (Algorithm 3):

- Initialize the sparse vector $\hat{\theta}$ and the iteration number i to 0, and the residual vector \mathbf{r} to \mathbf{y} .
- Correlate the residual \mathbf{r} against the columns of the reconstruction matrix \mathbf{A} .
- Update the solution by (a) scaling and adding the resultant correlation to the previous signal estimate; and (b) performing hard thresholding to keep the K largest values.
- Update the residual by subtracting $\mathbf{A}\theta^{(i+1)}$ from the measurement vector \mathbf{y} .

Algorithm 3 Iterative hard thresholding (IHT)

Input: \mathbf{A} the reconstruction matrix, \mathbf{y} the measurements, K the sparsity level, μ the step size, and the stopping criterion

Output: The sparse vector $\hat{\boldsymbol{\theta}}$

1: **Initialize:**

$$\boldsymbol{\theta}^{(0)} \leftarrow \mathbf{0}, \mathbf{r}^{(0)} \leftarrow \mathbf{y}, i \leftarrow 0$$

2: **while** stopping criterion not met **do**

3: $\mathbf{g}^{(i)} \leftarrow \mathbf{A}^T \mathbf{r}^{(i)}$

4: $\mathbf{z} \leftarrow \boldsymbol{\theta}^{(i)} + \mu \mathbf{g}^{(i)}$

5: $\boldsymbol{\theta}^{(i+1)} \leftarrow \text{Hthr}_K(\mathbf{z})$ ▷ keep the largest K entries of \mathbf{z}

6: $\mathbf{r}^{(i+1)} \leftarrow \mathbf{y} - \mathbf{A}\boldsymbol{\theta}^{(i+1)}$

7: $i \leftarrow i + 1$

8: **end while**

9: **return** $\hat{\boldsymbol{\theta}}$

In fact, the IHT operation can be described as

$$\boldsymbol{\theta}^{(i+1)} = \text{Hthr}_K(\boldsymbol{\theta}^{(i)} + \mu \mathbf{A}^T (\mathbf{y} - \mathbf{A}\boldsymbol{\theta}^{(i)})) \quad (2.29)$$

where μ is a step size, and i is the current iteration number. These steps are repeated until convergence. The parameter μ should be chosen carefully for rapid convergence. The algorithm can work better with adaptive stepsize (Blumensath and Davies, 2010). The IHT algorithm is easy to implement and is computationally efficient.

Combinatorial Approach

Algorithms under this approach pre-date CS, but they are relevant to sparse recovery. They were originally developed in the context of group testing to minimize the number of tests to be performed for solving sparse approximation problems. Popular combinatorial algorithms are random Fourier sampling, heavy hitters on steroids, chaining pursuits, and sparse sequential matching pursuit (Cormode, 2011). A specific pattern in the sensing matrix Φ is required (Φ needs to be sparse). For example, equal number of ones in each column of Φ but distributed randomly, which means that all the measurements in \mathbf{y} are obtained by combining the same number of samples of \mathbf{s} . These algorithms use two strategies called count-min and count-median. The steps (using count-min/median) are described as follows (Fig. (2.4)): To obtain the estimate $\hat{\mathbf{s}}_i$ of the i th sample \mathbf{s}_i of the input speech

- Identify all the measurements \mathbf{y}_j that have used \mathbf{s}_i with the help of the sensing matrix.
- Perform count-min/median strategy. For count-min, the estimate of the i th sample $\hat{\mathbf{s}}_i$ is the minimum value from the measurements identified in the previous step. For count-median, the median is computed instead of the minimum.

These algorithms are extremely fast and efficient compared to convex and greedy approaches.

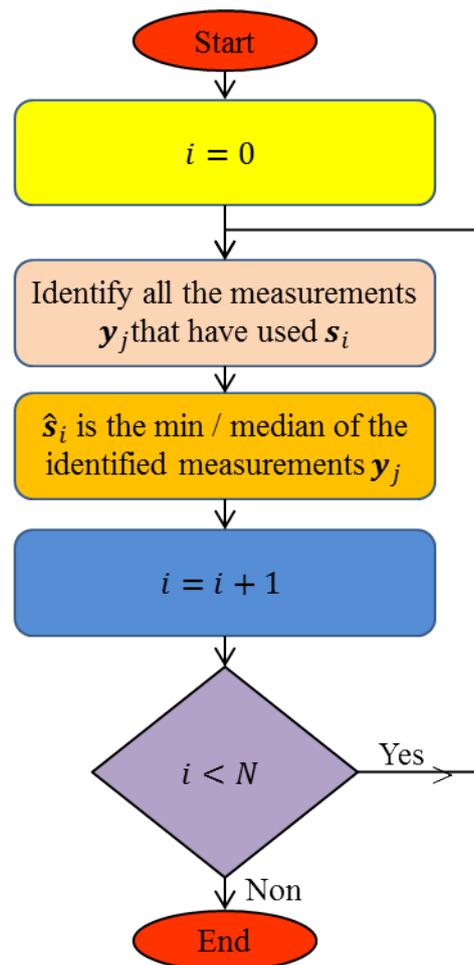


FIGURE 2.4: Combinatorial approach (using count min/median) steps

Non-Convex Approach

Algorithms under this approach search the solution with minimum ℓ_p norm where $0 < p < 1$, instead of ℓ_1 norm considered in convex approach. With this approach, much fewer measurements are required compared to convex approach. Examples are focal underdetermined system solution (FOCUSS), iteratively reweighted least squares (IRLS), *etc.* (Chartrand and Staneva, 2008).

Bayesian Approach

Unlike the previous approaches that consider deterministic input signals, Bayesian approach applies on signals with a known probability distribution. Hence, this approach seems to be of more practical interest. This approach considers the reconstruction as Bayesian inference problem. Maximum likelihood or maximum a posteriori can be used to estimate the coefficients of the input signal. Examples of algorithms under this category are belief propagation, sparse Bayesian learning using relevance vector machines, Bayesian compressive sensing, *etc.*

In principle, all the algorithms introduced in this section work reasonably well. The choice of an algorithm depends on the requirements in a specific situation, e.g., specific sensing matrix, the parameter values K , M , N , *etc.* A fixed (rough) choice can not be made as novel algorithms are being developed, and novel applications are being discovered continuously. Thus, an algorithm that is less competitive today might be the best choices in future under a new setting. Table 2.1 presents a summary of CS reconstruction approaches, which can help in selecting an approach that meets the system's requirement.

2.4 Applications of Compressed Sensing

When mathematicians developed CS, they had not any application in mind, they had only the general concept of subsampling as goal. Nevertheless, CS has afterwards been used in many areas such as medical imaging, astronomy, video processing, *etc.*

2.4.1 Compressive Imaging

Compressed sensing has been used in imaging to reduce the number of measurements, hence, power consumption, computational complexity, and storage space without sacrificing the spatial resolution. Some important applications include:

Single Pixel Camera

Single pixel camera was proposed by Duarte *et al.* (Duarte et al., 2008) in 2008. It consists of a digital micro-mirror devices (DMD) array that can either reflect light towards the sensor or reflect light away from it (on/off) using a pseudorandom pattern. Light received at photodiode is weighted average of many different pixels, whose combination gives a single pixel. The latter is then sampled by a low rate analog-to-digital converter (ADC) to generate the compressive measurements. These

TABLE 2.1: Compressed sensing reconstruction approaches

Approach	Attributes	Advantages	Constraints
Convex	<ul style="list-style-type: none"> - Global optimization method - Minimize the ℓ_1 norm 	<ul style="list-style-type: none"> - Robust to noise - Able to superresolve 	<ul style="list-style-type: none"> - Slower and complex - Difficult to implement for problems of larger size
Greedy	<ul style="list-style-type: none"> - Iterative method based on correlation 	<ul style="list-style-type: none"> - Faster, has low complexity, and robust to noise - Parallel versions are able to discard wrong entries selected in previous iterations 	<ul style="list-style-type: none"> - Requires prior knowledge about signal sparsity - Requires more measurements than convex approach - has convergence issues
Thresholding	<ul style="list-style-type: none"> - Uses some non linear thresholding criteria to select atoms 	<ul style="list-style-type: none"> - Faster and with low complexity - Able to add/discard multiple entries per iterations 	<ul style="list-style-type: none"> - Better performance requires adaptive step size, which increases complexity
Combinatorial	<ul style="list-style-type: none"> - Computes min / median of measurements identified as used a particular sample 	<ul style="list-style-type: none"> - Faster and simpler 	<ul style="list-style-type: none"> - Requires noiseless and specific pattern in measurements
Non-Convex	<ul style="list-style-type: none"> - Global optimization method - Minimize the ℓ_p norm with $0 < p < 1$ 	<ul style="list-style-type: none"> - Requires less measurements than ℓ_1 minimization - Number of measurements and error decrease with p 	<ul style="list-style-type: none"> - Slower and complex - difficult to implement for problems of larger size
Bayesian	<ul style="list-style-type: none"> - Consider recovery as Bayesian inference problem - Applicable for signals belonging to some known distribution 	<ul style="list-style-type: none"> - Faster and yields more sparser solution - Estimate signal parameters without user intervention 	<ul style="list-style-type: none"> - Has high computational cost

measurements can be easily stored or transmitted. The receiver can use CS reconstruction to recover the original scene. Single pixel camera can be used for colour images (hyperspectral camera) (Nagesh and Li, 2009b). Data captured by single pixel camera can be used for automatic detection and tracking objects (Elgammal, Harwood, and Davis, 2000). Terahertz imaging systems also achieve high speed

acquisition using single pixel detector (Chan et al., 2008). Improvements are achieved using dynamic single pixel camera for superresolution imaging (Zhao et al., 2017). Other examples of recent developments in this context are in multiresolution spectral imaging (Garcia, Correa, and Arguello, 2018), visible light imaging (Wu et al., 2018), superresolution imaging (Wei et al., 2019), biomedical imaging (Rousset et al., 2017), *etc.*

Medical Imaging

Medical imaging, particularly magnetic resonance imaging (MRI), has also benefited from CS. Magnetic resonance imaging is a common technology used for brain imaging, angiography (examination of blood vessels), dynamic heart imaging,... It is costly and time consuming (several minutes or hours depending on the task) when using the traditional Shannon sampling theorem. For instance, heart patients cannot be expected to hold their breath for too long time, and children are too impatient to sit still for more than about two minutes (Foucart and Rauhut, 2013). Since magnetic resonance images have sparsity properties in domains such as Fourier or wavelet basis, the introduction of CS has improved the image quality through reduction in the number of collected measurements. Magnetic resonance imaging using CS is an active research area. Many recent CS-based designs have been proposed in recent years, for example, (Tashan and Al-Azawi, 2018; Quan, Nguyen-Duc, and Jeong, 2018; Mathew and Paul, 2018; Sun et al., 2019; Yuan et al., 2019; Chen et al., 2019; Yang, Qin, and Wu, 2019; Mardani et al., 2019), *etc.*

Seismic Imaging

Seismic images are compressible in some transform domains, e.g., in curvelet basis (Herrmann and Hennenfent, 2008), time-frequency basis (Boashash, 2015, Chapter 14.6), *etc.* Moreover, seismic data is high-dimensional, incomplete, and very large. Compressed sensing has been exploited, for its dimensionality reduction property, to decrease massive data volume and high computational cost. Framework examples developed in this sense are (Sun et al., 2018; Zand et al., 2019; Cao, Hu, and Tang, 2019), *etc.*

Radar Imaging

For radar imaging, CS has been used in synthetic aperture radar (SAR) (Zhang and Hoorfar, 2015; Bi et al., 2018; Kang and Kim, 2019; Pu et al., 2019), inverse synthetic aperture radar (ISAR) (Zhang and Xing, 2019; Kang et al., 2017; Hou et al., 2017; Qiu, Zhou, and Fu, 2020), through the wall imaging radar (TWR) (Li and Burkholder,

2015; Wu et al., 2015; Wang et al., 2017; Tang, Bouzerdoum, and Phung, 2018; Wang et al., 2018), and ground penetrating radar imaging (GPR) (Tuncer and Gurbuz, 2012; Qu et al., 2015; Sun et al., 2016; Li et al., 2016). In all these techniques, CS exploits sparsity in these images to recover high resolution images from fewer measurements, simultaneously offering advantages like robustness and reduced storage.

Other imaging applications that have benefited from CS are parallel imaging (Peng et al., 2014), subwavelength imaging (Li, 2014), thermoacoustic imaging (Qin et al., 2015), microwave imaging (Guo and Abbosh, 2015), underwater imaging (Alqadah, 2016), *etc.*

2.4.2 Biomedical Applications

Apart from biomedical imaging, CS has been used with other biological signals such as electrocardiogram (ECG) (Bortolotti et al., 2018; Zhang et al., 2019), electroencephalographic (EEG) (Mammone et al., 2019; Li et al., 2019), neural signals (Xiong et al., 2018; Zhao et al., 2018), *etc.*, by exploiting their sparsity characteristic. Other biomedical applications are in microarrays (DNA, proteins,...) (Parvaresh et al., 2008; Mourad, Dawy, and Morcos, 2013).

2.4.3 Radar

In addition to radar imaging, CS can be applied in different radar frameworks. It can be used for monostatic, bistatic, and multistatic radar. Using CS, it is possible to detect, classify, and recognize target directly from incoherent measurements (Hu et al., 2017; Blacknell, 2016; Chen et al., 2017; Pieraccini and Miccinesi, 2019). The distance of an object and its speed can also be deduced (Foucart and Rauhut, 2013).

2.4.4 Pattern Recognition

Compressed sensing has been used in many recognition frameworks by exploiting the inherent sparsity in the pattern. For example, for face recognition (Nagesh and Li, 2009a; Zhang, Zhao, and Lei, 2012; Andrés et al., 2014), this exploits the sparsity of the expression changes in consideration of the whole image. Speech and speaker recognition (Sivaram et al., 2010; Gemmeke et al., 2010; Naseem, Togneri, and Bennamoun, 2010), gesture recognition (Akl and Valaee, 2010; Akl, Feng, and Valaee, 2011), gait recognition (Sivapalan et al., 2011), iris recognition (Bhateja et al., 2016), and computer vision (Wright et al., 2010) have all benefited from CS theory.

2.4.5 Video Processing

In video processing, CS has been used for video acquisition and reconstruction (Liu, Elezzabi, and Zhao, 2011), streaming video adaptive coding (Unde and Deepthi, 2019), high speed periodic video acquisition (Veeraraghavan, Reddy, and Raskar, 2010), 3D video acquisition using single pixel camera (Edgar et al., 2016), *etc.*

2.4.6 Manifolds Processing

The purpose of manifold models is to represent structure underlying the high dimensional data with the a small number of parameters. Compressed sensing has been applied to manifold-modeled data to achieve dimensionality reduction. Examples of the proposed frameworks are (Hegde, Wakin, and Baraniuk, 2007; Baraniuk and Wakin, 2007; Chen et al., 2010).

2.4.7 Micro and Nano-Electronics and VLSI Applications

Compressed sensing find also applications in areas of micro and nano-electronics and very large scale integration (VLSI). For example, it has been utilized for 3D visualization of the iron oxidation state in FeO/Fe₃O₄ core-shell nanocubes (Torruella et al., 2016), modeling the spatial variations of nanoscale integrated circuits by exploring sparsity due to correlated representation of spatial variations in frequency domain (Liao et al., 2016), low-cost silicon characterization of nanoscale integrated circuits (Zhang et al., 2011), testing vehicle for 3D TSV pre-bond and post-bond testing data (Huang et al., 2016), low complexity method for long term field measurement of insulator leakage current (Ghosh, Chatterjee, and Chakravor, 2016), *etc.*

2.4.8 Analog-To-Information Conversion

Because, in most applications, information content of the signal is much smaller than its bandwidth, sampling the whole signal is a wastage of resources. Therefore, CS has been used to solve this problem, by replacing the ADC by analog-to-information conversion (AIC). The latter utilizes random sampling for wideband signals to reduce the sampling rate; hence decreasing the consumption of hardware and software resources. Recent developments in this context are (Verhelst and Bahai, 2015; Wen, Tao, and Zhang, 2015; Wadhwa, Madhow, and Shanbhag, 2017; Pelissier and Studer, 2018; Xu, Zhang, and Kim, 2019), *etc.*

2.4.9 Speech and Sound Processing

Compressed sensing has exploited the sparsity of speech/audio and sounds for many purposes, e.g., speech coding, speech enhancement, audio security, sound localization, *etc.*

Speech Coding

Although studies have been conducted in speech coding using CS, they remain limited. In (Giacobello et al., 2012), a speech coding approach based on CS has been proposed. The authors have introduced sparsity in a linear prediction framework to improve the linear prediction coding (LPC) performance. (Al-Azawi and Gaze, 2018) has introduced a combined compression/encryption method for speech using compressed sampling. Contourlet transform increases the sparsity whereas the chaotic system generates the sensing matrix in the proposed approach. In the recently published paper by Ji *et al.* (Ji, Zhu, and Champagne, 2019), the authors have proposed a dictionary learning technique based on the recurrent neural network for CS of speech. A sequential linear prediction model and clustering have been used to generate codebooks for voiced and unvoiced speech. Additionally, a decision module has been designed to determine the appropriate dictionary for the recovery algorithm in the CS system. In (Haneche, Boudraa, and Ouahabi, 2018b) and (Haneche, Ouahabi, and Boudraa, 2019a), two novel speech coding methods based on the combination of CS with split vector quantization and split-multistage vector quantization, respectively, have been proposed. Wavelet matrix is used as the sparse basis, and convex optimization is considered for CS recovery. Details are provided in Chapter 3.

Speech Enhancement

For speech enhancement, few important CS-based methods have been proposed. Compressed sensing for speech enhancement seeks to estimate a sparse representation of the clean speech by reconstructing only the sparse components (speech) from the mixture (speech and noise). In (Sigg, Dikk, and Buhmann, 2012), a method called generative dictionary learning (GDL), based on learning speech and interferer signal models has been proposed. The speech and noise dictionaries were concatenated to provide one dictionary which was used to recover the clean speech from the noisy speech by the sparse coding. The authors presupposed that an appropriate training data for interferers was available to avoid using voice activity detection (VAD). Jančovič *et al.* (Jančovič, Zou, and Köküer, 2012) have presented a speech enhancement method based on both the sparse code shrinkage algorithm and the

use of multiple speech models. Gaussian mixture modelling was used in the training stage for clustering the speech signals in the independent component analysis-based transform domain into categories. For each category a Gaussian model was estimated and used for enhancement. Low *et al.* in (Low, Pham, and Venkatesh, 2013) performed the speech enhancement process on the spectral envelope of the short-time Fourier transform (STFT) of the noisy signal. They used CS to extract the sparse components (speech) from the mixture of sparse and non-sparse components (noise). In (Wu, Zhu, and Swamy, 2013; Wu, Zhu, and Swamy, 2014), Wu *et al.* established the theory of compressive sampling matching pursuit for speech enhancement in the case of time-domain noise (CoSaMPSE). They applied CoSaMPSE to denoise the overlapped frames in the discrete cosine transform domain. In (Ramdas, Mishra, and Gorthi, 2015), a design of an efficient medium bit-rate codec which can suppress the background noise has been presented. This approach uses analysis by synthesis to quantize CS measurement in the discrete cosine transform domain. Complementary joint sparse representations (CJSR) have been proposed in (Luo *et al.*, 2016). The authors aimed to make full use of the relationships among speech, noise, and mixture in sparse representation for speech enhancement. Both relationships were used to constrain the joint dictionary learning. Complementary advantages were achieved using weighting parameter based on Gini index. Wang *et al.* in (Wang *et al.*, 2016) considered speech enhancement as a problem of missing data imputation, and used CS to solve it in the frequency domain. To ensure sparsity, spectrograms of speech signals were used for learning an overcomplete dictionary by K-SVD algorithm. The enhancement process used also a mask designed based on the first formant detection method and a quasi-SNR criterion. Orthogonal matching pursuit (OMP) was used to find the K-sparse speech spectra. Another CS-based speech enhancement procedure is proposed in (Haneche, Boudraa, and Ouahabi, 2018a). In (Haneche, Boudraa, and Ouahabi, 2020), the information about the noise, in the measurement domain, is exploited besides sparse recovery to improve the performance of CS-based speech enhancement in low SNRs. More details can be found in Chapter 5.

Some other works applying CS to speech and audio signals focus mainly on audio processing (Stankovic and Brajovic, 2018), speech recovery from sub-Nyquist rate (Sreenivas and Kleijn, 2009), sparse LPC of speech in the residual domain (Giacobello *et al.*, 2010; Griffin *et al.*, 2011; Daniels and Rao, 2012; Giacobello *et al.*, 2012), speech recognition (Gemmeke *et al.*, 2010), source separation (Virtanen, 2007; Bao *et al.*, 2013), sparse decomposition of speech and audio signals (Christensen, Ostergaard, and Jensen, 2009), speech emotion recognition (Zong *et al.*, 2016), encryption (George, Augustine, and Pattathil, 2014; Cambareri *et al.*, 2015), declipping audio signals (Defraene *et al.*, 2013), voiced/nonvoiced detection in compressively sensed speech

signals (Abrol, Sharma, and Sao, 2015), packet loss recovery in audio multimedia (Ciaramella and Giunta, 2016). Similarly, CS for sound includes sound field reproduction (Lilis, Angelosante, and Giannakis, 2010), dynamic sound field measurement (Katzberg et al., 2018), ocean sound monitoring (Harris et al., 2016), sound sources localization (Lei et al., 2015), *etc.*

2.5 Compressed Sensing in Communication Systems

Compressed sensing is an attractive tool that can be used in the benefit of communication systems. Following, few interesting CS applications in communications. This list is not exhaustive, but represents a broad slice of current research areas.

2.5.1 Sparse Channel Estimation

Channel estimation concerns the estimation of the impulse response of the channel. This impulse response is used for the equalization of the received signal to mitigate the channel effect (e.g., intersymbol interference). Channel estimation is done by transmitting a known sequence, called pilot sequence, before transmitting the signal. Since the transmitted and the received sequences are known by the receiver, the only unknown is the channel impulse response. The latter may be estimated using sparse recovery because the channel impulse response may be assumed to be sparse (Kocic, Brady, and Stojanovic, 1995; Ariyavisitakul, Sollenberger, and Greenstein, 1997). This sparsity was exploited for channel estimation even before the apparition of CS (Cotter and Rao, 2002). Since the introduction of CS, this area has grown rapidly. It has been stated in (Bajwa et al., 2010) that one can use CS to reconstruct the channel impulse response with reduced energy consumption. Different scenarios have been investigated using CS. To cite some examples, (Berger et al., 2010a) found that an overcomplete dictionary may be used for better sparse representation. In (Berger et al., 2010b), CS have been used to channel estimation of multicarrier underwater acoustic communication systems. It has also been used to estimate the ultra-wideband channel in (Paredes, Arce, and Wang, 2007). The estimation of doubly selective channels using CS has been investigated in (Taubock et al., 2010). An optimized pilot placement for sparse channel estimation in orthogonal frequency division multiplexing (OFDM) systems is proposed in (Qi and Wu, 2011). Other recent frameworks on this area are (Chen, Liu, and Yuan, 2018; Li et al., 2018a; Akbarpour-Kasgari and Ardebilipour, 2019; Lee, Gong, and Chen, 2019; Zhang, Zhao, and Zhang, 2018; Uwaechia and Mahyuddin, 2019; Gómez-Cuba and Goldsmith, 2019; Zhang et al., 2019; Qian, Fu, and Sidiropoulos, 2019) , *etc.*

2.5.2 Wireless Sensor Networks

Another application area is in wireless sensor networks. In these networks a lot of sensor nodes (with communication capability) are used to measure some physical phenomenon (in the area of interest). Compressed sensing exploits the spatial and/or temporal correlation of the data sensed for fusion and collection. The objective here is not to lower the sampling rate, but rather to limit the amount of sensors so that less energy is spent on transmissions between sensors. The application of CS in wireless sensor network has been introduced by Bajwa et al. in (Bajwa et al., 2006). An interesting tutorial on this topic is (Haupt et al., 2008). Notable works on this topic include (Yang et al., 2013; Chen and Wu, 2015; Zayyani, Sari, and Korke, 2017; Li and Liang, 2018; Zhang and Wang, 2019; He et al., 2019; Zhou et al., 2019), *etc.*

2.5.3 Ultra Wideband Systems

Compressed sensing has been used in ultra wideband communication to reduce the high data rate of ADCs (i.e., acquisition rate of ultra wideband signals). Other issues like impulse radio detection (Wang et al., 2016), echo detection (Shi et al., 2008),... have also been addressed using CS. In (Gishkori and Leus, 2013) energy detection based on CS has been proposed for ultra-wideband pulse position modulation in multipath fading environments to reduce the sampling complexity at the receiver side. Authors in (Nguyen and Tran, 2016) have employed CS to address the problem of robust separation and extraction of multiple sources of radio-frequency interference from raw ultra-wideband radar signals in challenging bandwidth management environments. A deterministic signal-matched sparse measurement matrix for ultra wideband systems has been proposed in (Sharma, Gupta, and Bhatia, 2016; Sharma, Gupta, and Bhatia, 2019).

2.5.4 Cognitive Radio

In cognitive radio networks, unlicensed cognitive radio users (secondary users) are supposed to utilize licensed frequency bands during temporal vacancy of the users in the licensed systems (primary users) without causing harmful interference to them. It is, hence, necessary for the secondary users (or cognitive radio networks) to sense the radio-spectrum environment in order to find vacant frequency bands prior to communication. The latter task is very challenging using the conventional sampling theorem (Nyquist-Shannon) in the case of wide frequency band. By exploiting the sparsity in spectrum occupancy due to under-utilization of spectrum, CS can alleviate this problem (i.e., reduce the sampling rate). This is called compressed spectrum

sensing (Havary-Nassab, Hassan, and Valaee, 2010). It has been also extended to cooperative spectrum sensing, where observations of spectrum sensing at multiple secondary users are combined in order to achieve robustness against observation noise and/or fading effects (Bazerque and Giannakis, 2010; Zeng, Li, and Tian, 2011). Moreover, dynamic sampling rate adjustment has been considered in the compressed spectrum sensing schemes to cope with the time variant nature of the spectrum environment (Huang and Wang, 2012). Furthermore, the problem of primary user detection in cognitive radio has also been addressed (Başaran, Erköçük, and Çırpan, 2016). (Sharma et al., 2016) provides a survey on other applications of CS in cognitive radio. Some recent works in this topic are (Li et al., 2018b; Joneidi et al., 2018; Hamdaoui, Khalfi, and Guizani, 2018; Gong, Yang, and Zheng, 2019; Farrag, 2019; Eltabie, Abdelkader, and Ghuniem, 2019; Xu et al., 2019), *etc.*

2.5.5 Array Signal Processing

In addition to single measurement vector, CS has been also developed to address multiple measurement vector problem (Cotter et al., 2005; Chen and Huo, 2006), in other words, the problem of the recovery of a set of sparse signal vectors (sharing the same support). Direction-of-arrival estimation is one of the major problems in array signal processing. Compressed sensing can be applied to this problem due to the sparsity of the incoming signals in the angular domain. In (Gurbuz, McClellan, and Cevher, 2008), it has been demonstrated that using CS with all the array sensors except one, reduces the amount of data that must be communicated between the sensors. In (Kim, Lee, and Ye, 2012), an algorithm (called compressive MUSIC) has been proposed to solve the direction-of-arrival estimation problem which was formulated as a multiple measurement vector problem. Other recent works include (Hawes et al., 2017; Liu et al., 2018; Xu et al., 2018; Shi, 2019). Another application of CS to array signal processing is in designing maximally sparse array in order to achieve a minimum number of elements because the weight, the consumption, and the hardware/software complexity of the radiating device have a strong impact on the whole cost of the overall system (Oliveri and Massa, 2011; Prisco and D'Urso, 2012; Bencivenni et al., 2015; Bencivenni et al., 2016; Morabito, 2017). Compressed sensing has been also applied for arrays diagnosis (Migliore, 2011; Ince and Ögücü, 2016; Jiang et al., 2018; Li, Deng, and Migliore, 2018; Eltayeb, Al-Naffouri, and Heath, 2018; Salucci et al., 2018; Xiong et al., 2019; Palmeri, Isernia, and Morabito, 2019), *etc.*

2.5.6 Multiple Access Scheme

Multiple access means that multiple users send their data to a single receiver that needs to distinguish and reconstruct the data from these users. In these systems, the number of active users at one time can be typically far smaller than the number of potential users. Hence, taking advantage of the sparsity of active users, one can extract signals of active users by using algorithms of CS. The application of CS to multiple access schemes is similar to the conventional code-division multiple-access (CDMA) scheme, but with the signature codes shorter than the number of users (Shim and Song, 2012; He et al., 2018). Recent works in this area are (Alam and Zhang, 2018; Zhang et al., 2018; Du et al., 2018; Abebe and Kang, 2019; Vem et al., 2019; Yu, 2019), *etc.*

2.5.7 Network Tomography

Performance parameter estimation of internal links is often required for network management tasks such as congestion detection or traffic management. Network tomography allows the measurement of end-to-end to infer network internal links performance characteristics such as link packet loss rates (loss tomography) and delay (delay tomography). Packets are sent from source nodes and processed at receivers to get path performance measurements, then link characteristics are obtained by exploiting the dependence between links and corresponding paths (Fan, Li, and Zhang, 2018). An increase in the number of paths results in an increase in the number of probe packets injected into the network. In typical networks, only a limited number of links have large delays or high loss rates (Hayashi, Nagahara, and Tanaka, 2013). Hence, it is desirable to minimize the number of paths, which corresponds to the number of measurements, to avoid giving unnecessary load to the network. If we consider a vector composed by delays or loss rates of all links, we can reasonably assume that the vector will be sparse or compressible. Thus, the problem can come down to CS with a binary sensing matrix, which is determined by paths from source nodes to receiver nodes (Hayashi, Nagahara, and Tanaka, 2013). The authors in (Matsuda, Nagahara, and Hayashi, 2011) have proposed a loss tomography scheme based on CS with an $\ell_1 - \ell_2$ optimization for classifying communications links into higher and lower quality. In (Fan, Li, and Zhang, 2018), a loss tomography scheme based on CS is proposed using weighted ℓ_1 minimization. Authors in (Nakanishi et al., 2014) proposed a CS-based scheme for delay tomography. Other recent frameworks on delay estimation using CS are (Nakanishi et al., 2018; Wei et al., 2018; Kinsho et al., 2019), *etc.*

2.5.8 Multimedia Coding and Communication

Compressed sensing has great potential to be used in multimedia communication in applications such as wireless multimedia sensor networks. The latter are designed to transmit audio and video streams, images, and scalar data for many purposes such as video surveillance, object tracking, personal health care (telemedicine), theft control systems, traffic monitoring, agricultural monitoring (smart farms), industrial monitoring systems, *etc.* Wireless multimedia sensor networks produce an enormous amount of big multimedia data, which is a challenge for researchers additionally to energy consumption, reliability, signal detection and estimation, security- and privacy-related issues, *etc.* In recent years, various studies have focused on the use of CS in wireless multimedia sensor networks, and novel techniques for multimedia transmission are under investigation. A compressed sensing-based video encoding for transmission over wireless multimedia sensor networks has been studied in (Pudlewski, Prasanna, and Melodia, 2012) to reduce the encoder complexity and improve the resilience to lossy channel errors. In addition to the video encoder the system consists of a distributed rate controller and an adaptive parity channel encoding scheme to encounter the errors of the binary symmetric channel and to provide to the receiver high-quality videos. In (Pudlewski and Melodia, 2013b), improvements to the latter compressive video streaming method are achieved using novel rate-energy-distortion analysis. The model allows comparing the received video quality, computation time, and energy consumption per frame of different wireless streaming systems, and can be used to determine the optimal allocation of encoded video rate and channel encoding rate for a given available energy budget. Pudlewski and Melodia in (Pudlewski and Melodia, 2013a) provide a tutorial on encoding and wireless transmission of compressively sampled videos. In (Qi et al., 2015), a hybrid security and CS-based scheme for multimedia sensor data gathering has been proposed to reduce data volume and keep security. The proposed scheme has light security mechanism and thus decreases the complexity and energy consumption of system. Chen *et al.* in (Chen et al., 2018) have proposed an efficient and robust image coding and transmission based on scrambled block CS. By using scrambled block CS, progressive non-uniform quantization, and progressive non-local low-rank reconstruction, the transmission can achieve simple encoding, good rate-distortion performance, and robustness against packet loss. Bavarva *et al.* in (Bavarva, Jani, and Ghetiya, 2018) have improved the performance of wireless multimedia sensor networks, in terms of energy consumption and image quality in Rayleigh fading environment, by exploiting the multiple input multiple output (MIMO) properties and using CS to acquire fewer measurements of the original image. Authors in (Li et al., 2020) have proposed an image communication system exploiting CS for

internet of things (IoT) monitoring applications to satisfy the requirements of sensor nodes for low computational complexity, low energy consumption, and low storage overhead. The transmission security has also been enhanced through the proposed CS model and using quantization and diffusion operations into the system.

Future mobile systems plan to provide high data rates (more than 100 Gbps). In addition to the increasing demand for high-quality multimedia, users expect a high speed, high reliability, and real-time communication at anytime and anywhere. In our paper (Haneche, Ouahabi, and Boudraa, 2019b), a new mobile communication system design for Rayleigh environments based on CS-source coding has been proposed to increase transmission speed, robustness, and security in order to meet the requirements of 5G mobile systems that know an exponentially increasing data amount over time. The end-to-end communication scheme relies on the use of CS for speech coding instead of the supported coding standards in actual mobile communication systems, performing multicarrier modulation, and using channel coding and receive antenna diversity. This entails mitigating the channel effects, optimizing spectral efficiency, and ensuring robustness and reliability while increasing the baud rate. The proposed system shows a simplified design, and allows reducing bit rates and processing load compared to actual communication systems. The recovered speech has good quality and fair intelligibility scores. In addition, CS allows secure communications without additional costs. More details are provided in Chapter 4.

2.5.9 Information Security

The applications of CS in the field of information security have captured a great deal of researcher's attention in the past decade. The basic idea is to enable measurement matrix to be a key known by the encoder and the decoder. A review on CS in information security field is provided by Zhang *et al.* in (Zhang *et al.*, 2016). Some other recent works include (Peng *et al.*, 2017; Qing, Guangyao, and Xiaomei, 2018; Gan, Xiao, and Zhao, 2018; Li *et al.*, 2019; Mangia *et al.*, 2019; Chen *et al.*, 2019; Wang *et al.*, 2019), *etc.*

2.6 Conclusion

We have presented in this chapter the theoretic fundamentals of CS, and we have reviewed the major applications of this paradigm including the applications in communication systems. The subsequent chapters detail our contributions. The next chapter presents the application of CS for speech compression.

Chapter 3

Compressed Sensing-Based Speech Codec

3.1 Introduction

With the high demand for the use of multimedia, including speech, the volume of data circulating in the network has grown exponentially; hence the need to find better methods for compression (Ferroukhi et al., 2019). We focus, in this chapter, on the use of compressed sensing (CS) for speech coding. Although CS permits to collect only few measurements that allow future reconstruction, it is not considered as a compression in the strict-sense of the word; a quantization is needed to provide binary data that can be processed by digital systems. In this chapter, two proposed speech codecs are presented. The first is based on split vector quantization (SVQ) of CS measurements (Haneche, Boudraa, and Ouahabi, 2018b), and the second uses split-multistage vector quantization (S-MSVQ) (Haneche, Ouahabi, and Boudraa, 2019a). In addition, an overview of quantization methods; namely scalar (SQ), vector quantization (VQ), and multistage vector quantization (MSVQ), is provided with new results for speech coding.

3.2 Motivation

Current speech coding systems perform compression after the acquisition of signals at the Nyquist rate. This causes unnecessary processing burden and resources consumption as most signals can be represented by few significant coefficients in the appropriate basis (sparsity). Additionally, current speech coding standards, such as G.722.2, require high computational complexity. Taking the advantage of sparsity of speech in the wavelet basis, our idea is to exploit the dimensionality reduction property of CS to design a new speech codec that reduce the complexity

and the consumption of physical resources (e.g., energy, memory, *etc.*). The proposed

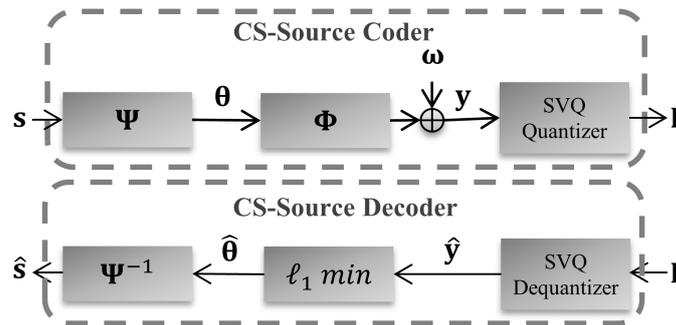


FIGURE 3.1: The proposed compressed sensing-based speech codec using split vector quantization (SVQ). (This codec performs SVQ quantization of the observation vector \mathbf{y} . Note that ω is the measurement noise due to compressed sensing.)

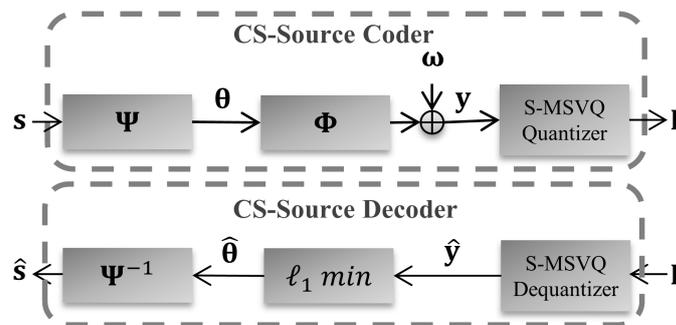


FIGURE 3.2: The proposed compressed sensing-based speech codec using split-multistage vector quantization (S-MSVQ). (This codec performs S-MSVQ quantization of the observation vector \mathbf{y} . Note that ω is the measurement noise due to compressed sensing.)

CS-speech encoders include, as depicted in Figs. 3.1 and 3.2, a CS module and a quantization module. The compressed sensing-speech decoders, in the other side, perform dequantization and CS reconstruction.

Indeed, quantizing CS measurements for the purpose of compression has been investigated before, but these studies have been conducted on other signals than speech. In (Baig, Lai, and Lewis, 2010), the effects of quantization on CS video are studied. Quantized spectral CS is examined in (Fu and Chi, 2018). Scalar quantization of CS measurements for real-time energy-efficient electrocardiogram (ECG) compression on wireless body sensor nodes is investigated in (Mamaghanian et al., 2011). In (Laska and Baraniuk, 2012), a CS system with SQ is studied for noisy input signals. In (Craven et al., 2014), non-uniform quantization of CS measurement for ECG signals is proposed. Optimal quantization of random CS measurements is studied in (Sun and Goyal, 2009). Differential pulse code modulation (DPCM) for quantized CS of

images is proposed in (Mun and Fowler, 2012). In (Krahmer, Saab, and Yilmaz, 2014), sigma-delta quantization of CS measurements is proposed. Analysis-by-Synthesis quantization for CS measurements is studied in (Shirazinia, Chatterjee, and Skoglund, 2013). In (Dai and Milenkovic, 2011), the average distortion introduced by SQ, VQ, and entropy coded quantization of CS measurements is studied. Performance for vector quantized CS is investigated in (Shirazinia, Chatterjee, and Skoglund, 2012). Joint source-channel coding of CS measurements using VQ and MSVQ is proposed in (Shirazinia, Chatterjee, and Skoglund, 2014). Moreover, studies on speech coding using CS remain limited (see section 2.4.9). In this chapter, we present new results for SQ, VQ, and MSVQ of CS measurements for speech coding. In addition, we propose two alternative speech codecs based on SVQ and S-MSVQ of CS measurements. These two quantizations methods have not been used for quantizing CS measurements in prior works.

3.3 Overview of Quantization Methods

The process that allows presenting an infinite set of numbers \mathbf{y} to a smaller set of quantized values $Q(\mathbf{y})$ is called quantization. The quantizer encoder maps samples in the same region to an index. Based on this index, the quantizer decoder performs the inverse quantization by allocating the codeword that corresponds to this index. The quantization is an irreversible mapping process; consequently, an error is introduced. Quantization methods used in this work are introduced next.

3.3.1 Scalar Quantization

The principle of scalar quantization (SQ) (Fig. 3.3) is to allocate a code to every sample. Let $\mathbf{y} \in \mathbb{R}^M$ be the measurement vector to be quantized, SQ is the process that maps each CS-sample $\mathbf{y}(i)$ to a codeword $\mathbf{q}(j)$ from a codebook $\mathcal{C} = \{\mathbf{q}(j)/j = 1, \dots, L\}$, where L is the number of quantization levels. In this thesis optimized SQ is considered by using Lloyd (Lloyd, 1982) algorithm to construct the codebook.

3.3.2 Vector Quantization

Vector quantization (VQ) (Fig. 3.4) is known as an optimal quantization tool when the coding rate is low. It deals with quantizing a group of samples (or vectors) rather than individual samples; hence the name. A VQ encoder of dimension dim (note that SQ is a special case of VQ, obtained if $dim = 1$) and length $L = 2^B$, where B is the number of bits available, considers vectors \mathbf{u} of dimension dim from the observation

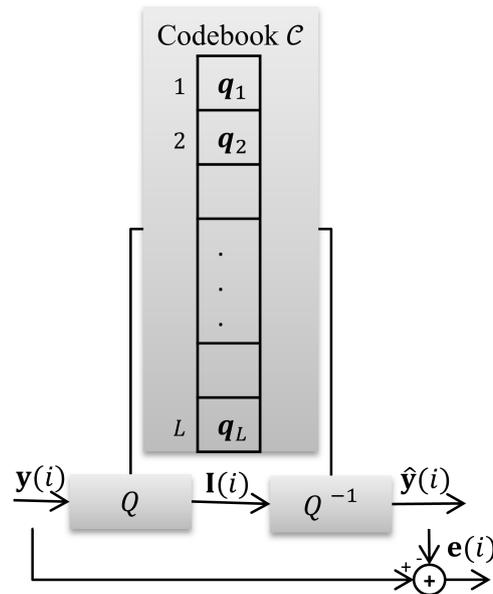


FIGURE 3.3: Diagram of scalar quantization

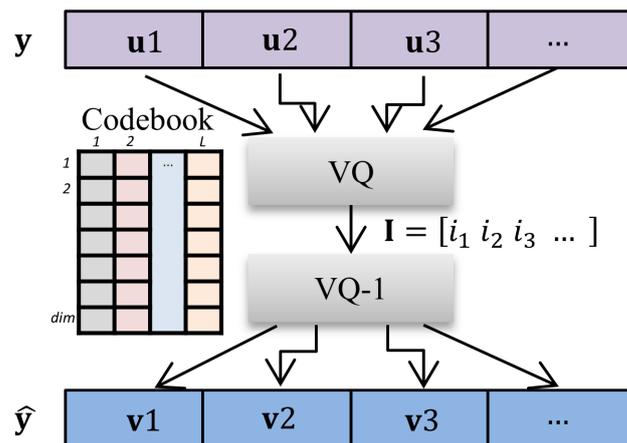


FIGURE 3.4: Diagram of vector quantization

vector \mathbf{y} , and affects them indexes from a codebook \mathcal{C} of dimension $dim \times L$ that contains L codevectors of dimension dim . The codevectors \mathbf{v} are the representation vectors of each Voronoi region¹. An iterative algorithm called Linde, Buzo and Gray (LBG) (Linde, Buzo, and Gray, 1980) is used for codebooks design. For applications requiring high bit-rates, the computational complexity and memory requirements increases exponentially. There exists various VQ techniques to deal with the later issues, e.g., multistage VQ, split VQ, and split-multistage VQ.

¹The space is partitioned to L classes \mathbf{R}_i , called Voronoi regions, such that $\mathbf{R}_i = \{\mathbf{u}/Q(\mathbf{u}) = \mathbf{v}\}$, \mathbf{v} is the centroid of the region \mathbf{R}_i

3.4 Speech Codec Based on Multistage Vector Quantization of Compressed Sensing Measurements

Multistage vector quantization (MSVQ) is an evolution of the basic VQ technique, also known as multistep, residual or cascaded VQ. It is a cascaded VQ encoders where the output of the VQ of a stage is the input of the next stage. It keeps the features of VQ technique while decreasing the computational complexity and memory requirements and enhancing the quality. In MSVQ, every stage has its own codebook and the bits available B are shared between the st stages: In the i th stage, B_i bits are used to create a codebook of dimension $dim \times L_i$, where $L_i = 2^{B_i}$ and $B = \sum_{i=1}^{st} B_i$. The codebook of the first stage is created using the training sequence as input, and the codebook of each other stage is created using the quantization error of its previous stage. As for VQ, LBG algorithm is used to create the codebooks.

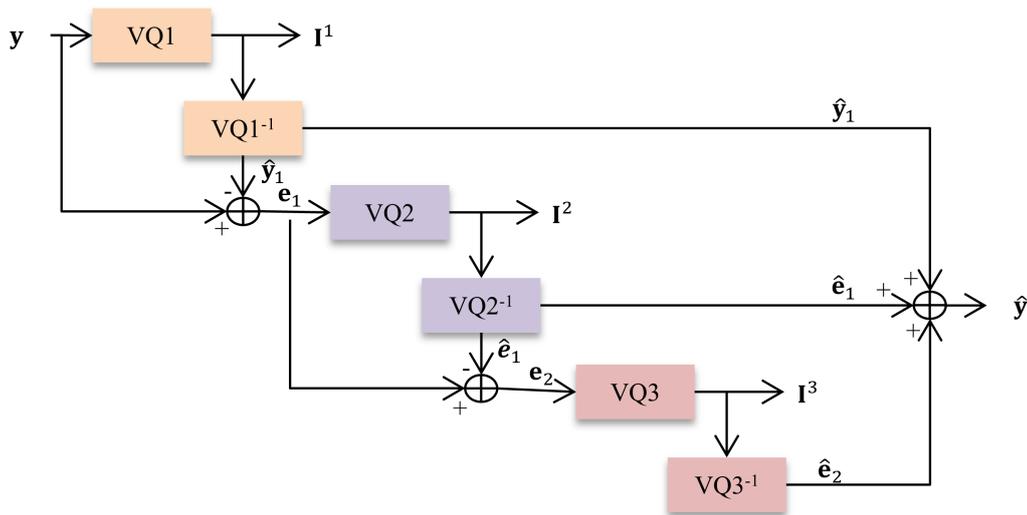


FIGURE 3.5: Diagram of three stages MSVQ (note that at stage t ($1 \leq t \leq 3$) the vector of indexes is $\mathbf{I}^t = [i_1^t i_2^t \dots]$).

Let y be the signal to be quantized, and assuming the number of stages $st = 3$. Fig. 3.5 illustrates the MSVQ encoding and decoding. Note here that the input of the VQ of the first stage is y , the input of VQ of the second stage is the quantization error of the previous stage $e_1 = y - \hat{y}_1$, and the input of VQ of the third stage is the quantization error of stage two $e_2 = e_1 - \hat{e}_1$. The MSVQ encoder provides three indexes corresponding to the observation vector y , the quantization error of the first stage e_1 , and the quantization error of the second stage e_2 . The MSVQ decoder uses the codebook of every stage to find the corresponding codevectors of the indexes, and provides the quantized vectors of the three stages \hat{y}_1 , \hat{e}_1 , and \hat{e}_2 . The quantized observation vector is then $\hat{y} = \hat{y}_1 + \hat{e}_1 + \hat{e}_2$.

3.5 Speech Codec Based on Split Vector Quantization of Compressed Sensing Measurements

This speech codec (see Fig. 3.1) has been proposed in (Haneche, Boudraa, and Ouahabi, 2018b). Split vector quantization (SVQ) is another idea used to reduce complexity/memory requirements of VQ. It consists of splitting vectors into smaller dimension subvectors. The bits allocated for quantization are, consequently, shared between the splits. In SVQ a separate codebook is created for every split. For sp -split VQ, the vectors (of dimension dim) of the training sequence are split into sp subvectors of smaller dimension dim_i to create the sp codebooks of dimensions $dim_i \times B_i$, where the bits available for quantization $B = \sum_{i=1}^{sp} B_i$ and $dim = \sum_{i=1}^{sp} dim_i$. Therefore, less computational complexity and memory are used, as they depend on the vector dimension and the number of bits available. LBG algorithm is used for the codebooks creation. Note that SVQ acts as SQ if the number of splits $sp = dim$.

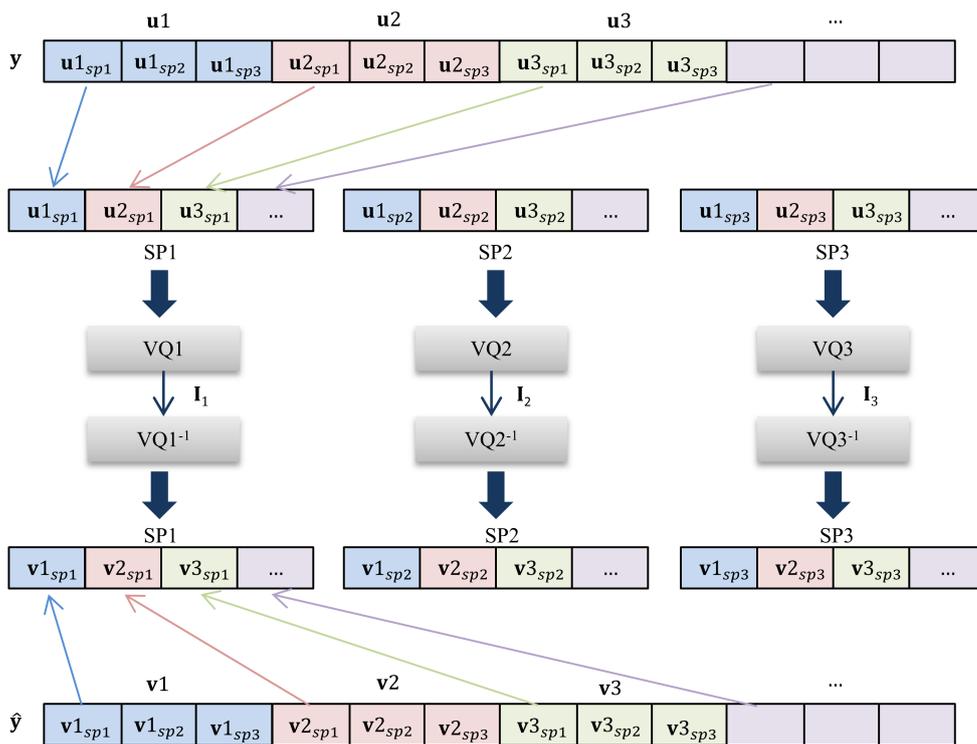


FIGURE 3.6: Diagram of three splits vector quantizer (note that for a split p ($1 \leq p \leq 3$) the vector of indexes is $I_p = [ip_1 ip_2 \dots]$).

Let's assume that the number of splits $sp = 3$. To perform SVQ (Fig. 3.6) on the observation vector y , we firstly consider vectors u of dimension dim (as in VQ), then these vectors are partitioned to subvectors u_{sp1} , u_{sp2} , and u_{sp3} of dimensions dim_1 , dim_2 , and dim_3 respectively. The subvectors of each split are quantized using

the proper codebook to get three indexes for every vector \mathbf{u} of \mathbf{y} . The SVQ decoder uses the appropriate codebooks to find the quantized subvectors \mathbf{v}_{sp1} , \mathbf{v}_{sp2} , and \mathbf{v}_{sp3} corresponding to the indexes. The concatenation of the three vectors results in the quantized vector $\hat{\mathbf{y}}$.

3.6 Speech Codec Based on Split-Multistage Vector Quantization of Compressed Sensing Measurements

This speech codec (see Fig. 3.2) has been proposed in (Haneche, Ouahabi, and Boudraa, 2019a). Split-multistage vector quantization (S-MSVQ) is a hybrid quantization method developed by combining SVQ and MSVQ. It is a cascaded SVQ encoders. In other words, it performs SVQ on the observation vector \mathbf{y} at the 1st stage, and on the quantization error for the other stages. The concatenation of the resultant quantized splits (of the same stage) and the sum of the quantized vectors of the different stages are performed to obtain the quantized observation vector $\hat{\mathbf{y}}$. Summing the quantized errors of each stage reduces the distortion. Decreasing the codebooks dimensions (by splitting and sharing the available bits between the splits and stages) results in using less complexity and memory. For each split of every stage, a separate sub-codebook is used. Split-multistage vector quantization divides the bits available B among the st stages and sp splits by allocating B_i^j bits for the i th split of stage j , where $B = \sum_{j=1}^{st} \sum_{i=1}^{sp} B_i^j$. In fact, B is shared between the st stages such that B^j bits is used for the stage j , where $B = \sum_{j=1}^{st} B^j$, and for each stage j the bits available B^j are shared between the sp splits such that $B^j = \sum_{i=1}^{sp} B_i^j$.

Fig. 3.7 illustrates the codebooks creation process for a 3-parts (splits) and 3-stages S-MSVQ. At the 1st stage, the vectors (of dimension dim) of the training sequence are partitioned into 3 subvectors of smaller dimension dim_i , where $dim = \sum_{i=1}^3 dim_i$, and used as inputs in LBG algorithm to generate the three sub-codebooks of dimensions $dim_i \times L_i^1$ ($L_i^1 = 2^{B_i^1}$, where B_i^1 is the numbers of bits available for stage 1 and split i). For the 2nd and the 3rd stages, LBG algorithm is performed on the quantization error of the 1st and the 2nd stages, respectively. As a result, 9 sub-codebooks are created.

Assuming the number of stages $st = 3$ and the number of splits $sp = 3$, Fig. 3.8 presents the quantization of \mathbf{y} by S-MSVQ using the obtained codebooks. To perform S-MSVQ on the observation vector \mathbf{y} , we firstly consider vectors \mathbf{u} of dimension dim (as in VQ), that we partition to subvectors \mathbf{u}_{sp1} , \mathbf{u}_{sp2} , and \mathbf{u}_{sp3} of dimensions dim_1 , dim_2 , and dim_3 respectively (as in SVQ). Using the appropriate sub-codebook, VQ is performed giving, hence, the quantized subvectors of the 1st stage $\hat{\mathbf{u}}_{sp1}$, $\hat{\mathbf{u}}_{sp2}$,

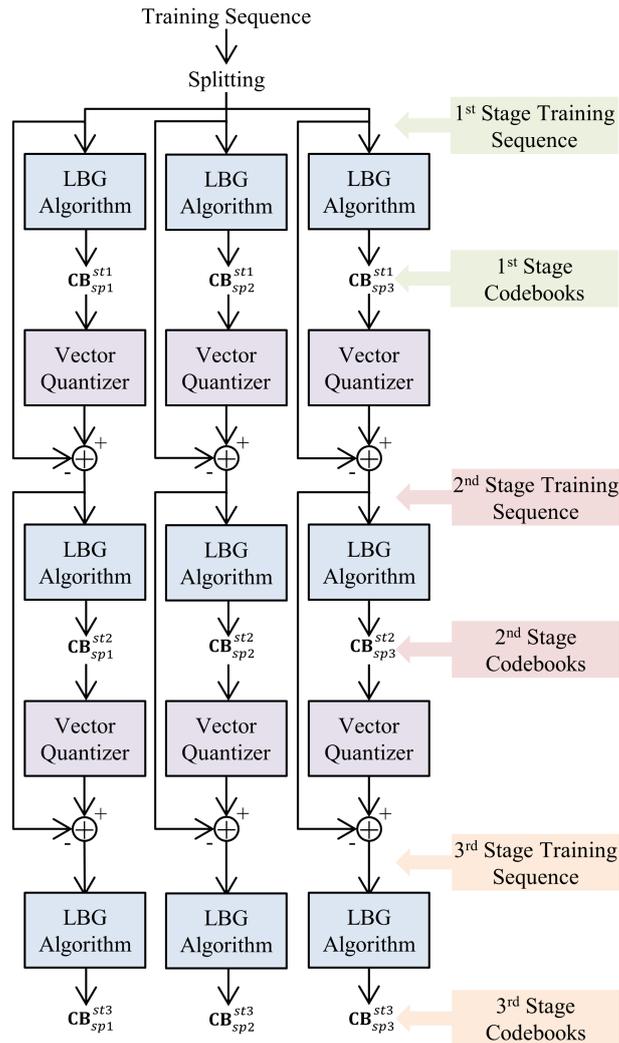


FIGURE 3.7: Codebooks Generation for three-parts three-stages split-multistage vector quantizer (S-MSVQ). The training sequence is partitioned to three splits. LBG algorithm is performed on the three splits at the 1st stage, and on the quantization error (of the previous stage) for the 2nd and the 3rd stages. As a result, nine sub-codebooks are created

and $\hat{\mathbf{u}}_{sp3}$. For the 2nd stage, VQ is performed on the quantization error of the three subvectors at the 1st stage \mathbf{e}_{sp1} , \mathbf{e}_{sp2} , and \mathbf{e}_{sp3} , where $\mathbf{e}_{sp1} = \mathbf{u}_{sp1} - \hat{\mathbf{u}}_{sp1}$, which results in the quantized error subvectors $\hat{\mathbf{e}}_{sp1}$, $\hat{\mathbf{e}}_{sp2}$, and $\hat{\mathbf{e}}_{sp3}$. The quantization errors of the 2nd stage \mathbf{w}_{sp1} , \mathbf{w}_{sp2} , and \mathbf{w}_{sp3} , where $\mathbf{w}_{sp1} = \mathbf{e}_{sp1} - \hat{\mathbf{e}}_{sp1}$, is used as input of VQ at the 3rd stage VQ encoder, giving the quantized error subvectors $\hat{\mathbf{w}}_{sp1}$, $\hat{\mathbf{w}}_{sp2}$, and $\hat{\mathbf{w}}_{sp3}$. The summation of the quantized splits of the three stages results in the quantized subvectors \mathbf{v}_{sp1} , \mathbf{v}_{sp2} , and \mathbf{v}_{sp3} (e.g., for the 1st split, $\mathbf{v}_{sp1} = \hat{\mathbf{u}}_{sp1} + \hat{\mathbf{e}}_{sp1} + \hat{\mathbf{w}}_{sp1}$). The concatenation of the three quantized subvectors results in the quantized vectors \mathbf{v} that constitutes the quantized observation vector $\hat{\mathbf{y}}$.

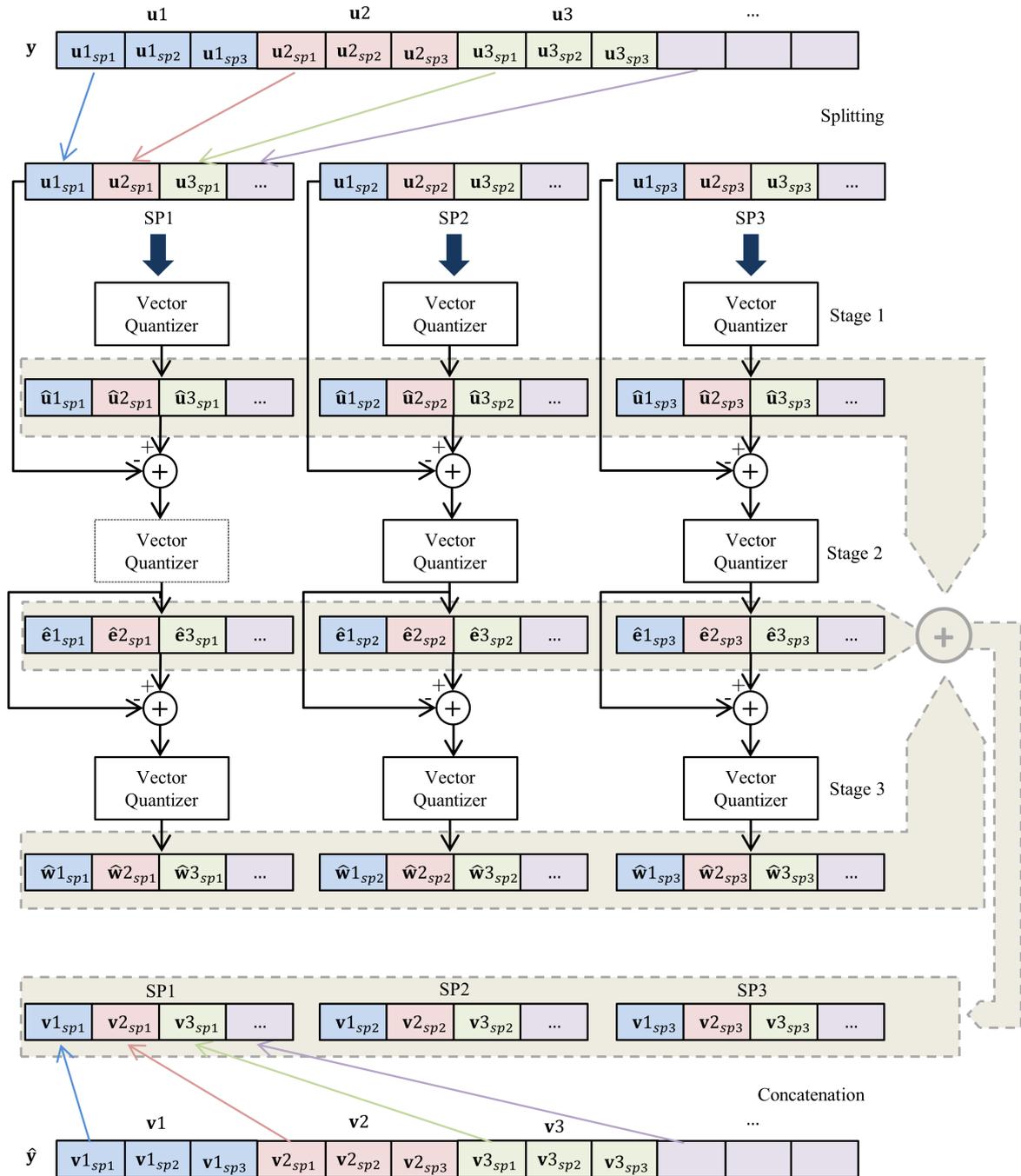


FIGURE 3.8: Diagram of three-splits three-stages split-multistage vector quantizer (S-MSVQ). The quantized observation vector \hat{y} is obtained by summing the quantized subvectors of the three stages and concatenating the resultant splits

3.7 Experiments

The purpose of this work is to compress speech using less memory and processing burden. To this end, we use CS to simultaneously acquire and compress signals, which means that non-significant data are neither acquired nor processed. Hence,

fewer samples are collected, and need to be quantized to allow its processing by digital systems.

The proposed compression methods (i.e., combining CS with SVQ or S-MSVQ) presented in Figs. 3.1 and 3.2 are applied for speech compression, and the results are compared with those obtained when other quantization techniques; namely SQ, VQ and MSVQ are used. Fig. 3.9 shows the simulation system: The speech signal is acquired by CS and then quantized using different methods. After that, inverse quantization and CS reconstruction are performed to compare the effects of the studied quantization methods.

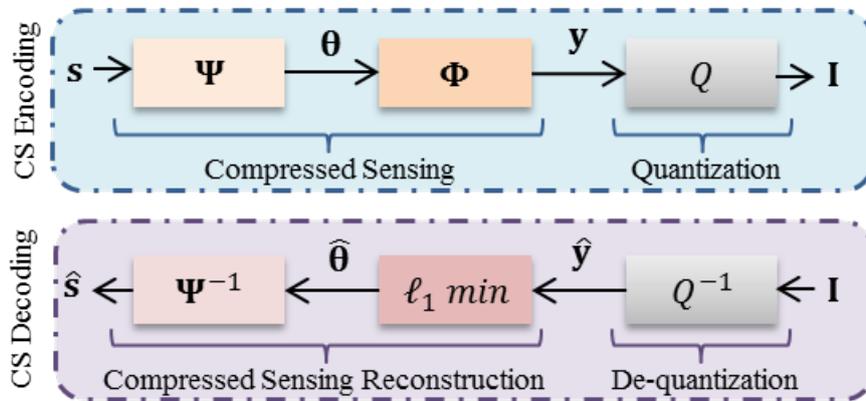


FIGURE 3.9: Simulation system.

The simulation system is showed in Fig. 3.9. The steps consist of:

- Acquiring compressed measurements using CS as follows
 1. Sparsify the speech segments \mathbf{s} by the projection on the wavelet basis Ψ and performing hard thresholding; and
 2. perform acquisition by the multiplication with a random Gaussian matrix Φ , which provide the CS measurements (the observation vector).
- Applying the quantization techniques Q on CS measurements \mathbf{y} , giving thus indexes \mathbf{I} from the appropriate codebooks.
- De-quantizing using the quantizer decoders of the studied quantization methods to find the corresponding codewords/codevectors that constitute the quantized observation vector $\hat{\mathbf{y}}$.
- Performing CS-reconstruction using ℓ_1 minimization to recover the sparse vector $\hat{\boldsymbol{\theta}}$ of length N from the quantized observation vector $\hat{\mathbf{y}}$ (of length M). The speech segments $\hat{\mathbf{s}}$ are obtained by inverse transformation (Ψ^{-1}).

3.7.1 Simulation Setup

Speech signals used in this work come from TIMIT database (Garofolo et al., 1993). Some sentences were considered for training, and others for testing. These signals are sampled at $f_s = 16$ kHz. For CS, the speech was processed by segment using a Hamming window of length $T_f = 25$ ms (which correspond to $N = f_s \times T_f = 400$ samples per each speech segment \mathbf{s}). Different compression ratios were considered when performing CS. The compression ratio C is described as

$$C = M/N \quad (3.1)$$

where M and N are the lengths of the observation vector \mathbf{y} , and the speech segment \mathbf{s} , respectively.

Codebooks for vector quantizations were created using a training database composed of CS measurements (observation vectors) obtained by performing CS on the training speech signals at different compression ratios C and various number of bits per vector. Thus, we have a different training database for each C and every number of bits per vector. Note that 50 utterances were considered for training, and 20 utterances, different from the latter, were used for test (of both genders).

3.7.2 Performance Measures

The performances of the proposed speech compression techniques are evaluated in terms of:

1. Mean-squared error (MSE) which is described as

$$MSE = \frac{1}{N} \sum_{i=1}^N (\mathbf{s}(i) - \hat{\mathbf{s}}(i))^2 \quad (3.2)$$

where \mathbf{s} and $\hat{\mathbf{s}}$ are the original and the reconstructed (from the quantized CS measurements) speech signals, respectively.

2. The wideband extension of the perceptual evaluation of speech quality (WB-PESQ). The WB-PESQ is an objective speech quality measure, which has a very close correlation with the mean opinion score (MOS). It is a standardized algorithm defined in ITU-T recommendation P.862.2 (P.862.2, 2007). The range of PESQ lies within -0.5 (bad) to 4.5 (excellent).
3. Speech intelligibility, which measures the amount of speech items that are recognized correctly. A frequently used objective measure is the speech intelligibility

index (SII). The latter is standardized by ANSI standard 3.05 (1997). Its range lies from 0 to 1. Speech of *good* intelligibility has a $SII > 0.75$, and speech of *poor* intelligibility has a $SII < 0.45$. In this work, coherence SII (Kates and Arehart, 2005) is used to measure the intelligibility performance. Fig. 3.10 gives more details on the PESQ and SII scales.

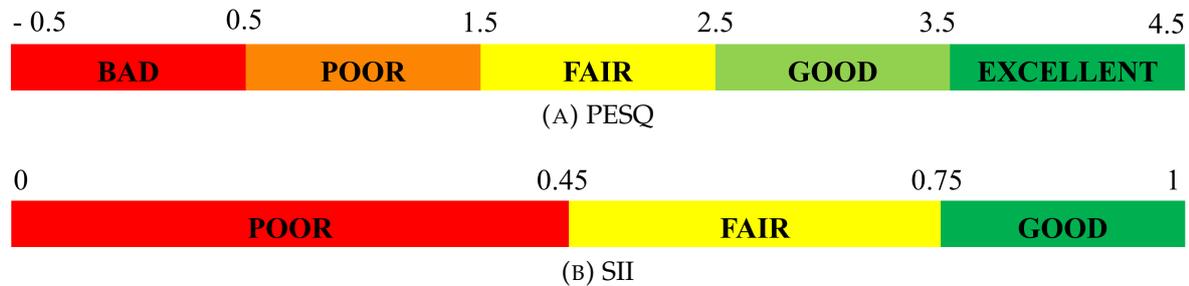


FIGURE 3.10: The PESQ and SII scales

The execution time of quantization/dequantization is also used for performance comparison.

3.7.3 Results and Discussion

We provide numerical results to analyse the effects of the different quantization methods on the performance of the recovered signal when speech compression is the purpose. The results presented here are averaged over all the speech signals considered for test.

Evaluation of CS-Codec using Scalar Quantization

The first experiment was carried out for scalar quantization of CS measurements using different compression ratios (Eq. (3.1)) and numbers of bits per sample ($nbps$). Fig. 3.11 presents the average PESQ, MSE, and SII scores of the recovered speech from quantized CS measurement using SQ, for various values of C and for a number of bits per sample ranging from 1 to 10. One can note (from Fig. 3.11(A)) the rising of PESQ score with the rising of C and $nbps$ for $C \leq 50\%$ and $nbps \leq 4$ bits/sample, it is almost constant after these values. We also notice that one can get *good* speech quality using only 2 bits/sample with $C = 10\%$ (e.g., PESQ= 2.54 for SQ), or 1 bit/sample with the other compression ratios (e.g., PESQ= 2.56 for SQ with $C = 20\%$).

It is also noticed from Fig. 3.11(B) that the average MSE value decreases with the rising of the compression ratio (C) and the number of bits per sample ($nbps$), and it remains constant when $C \geq 60\%$ and $nbps \geq 4$. Low MSE values are obtained for all C and $nbps$.

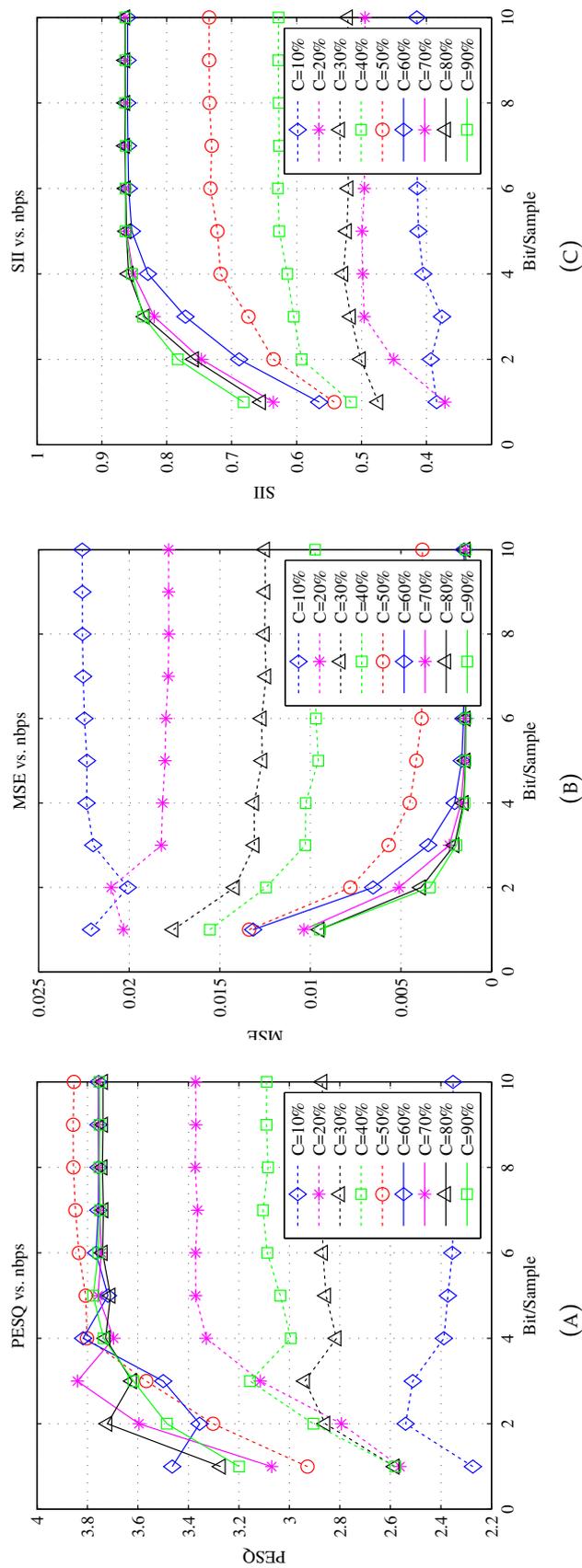


FIGURE 3.11: Performance of speech compression through scalar quantization of CS measurements for an *nbps* varying from 1 to 10 and for various compression ratios (C).

The same remark is depicted from Fig. 3.11(C) where the SII value increases until $C = 60\%$ and $nbps = 4$. It is also apparent from this figure that using only 2 bits per sample, *fair* average SII scores are obtained for $20\% \leq C \leq 60\%$, and *good* SII scores are achieved for $C \geq 70\%$. Using 1 bit per sample, *fair* average SII scores are obtained for $C \geq 30\%$.

From this experiment, there is evidence that a trade-off between the PESQ score and the bitrate is achieved when $C = 20\%$ and $nbps = 4$: The average PESQ score is 3.33 for a bitrate of 12.8 kbit/s. The bitrate for SQ is computed as

$$BR = nbps \times M / T_f \quad (3.3)$$

where $nbps$ is the number of bits per sample, T_f is the frame duration which is equal to 25 ms. and M is the observation vector length.

$$M = C \times N \quad (3.4)$$

where C is the compression ratio, and $N = T_f \times f_s$ is the speech segment length; f_s is the sampling frequency. By substitution, $M = C \times T_f \times f_s$, and the bit-rate expression can be simplified to

$$BR = C \times f_s \times nbps \quad (3.5)$$

The BR changes with the compression ratio (C) and the number of bits per sample ($nbps$).

To draw attention to the effect of SQ, another experiment was done by performing CS, then reconstruction without the quantization/dequantization steps. The average performance scores of the recovered speech are shown in Fig. 3.12. The later figure highlights also the comparison with the performance of SQ of CS measurements when $nbps = 4$. This test reveals that the SQ error does not significantly affect the CS reconstruction. In other words, the performance assessments (average PESQ, MSE, and SII scores) of the recovered speech from the quantized observation vectors are similar to those of the recovered speech from the analogue observation vectors.

Evaluation of CS-Codecs using Vector Quantizations

To evaluate speech coding using CS with VQ, SVQ, MSVQ, and S-MSVQ, other experiments proceeded following two cases. In the first case, we fixed the bitrate

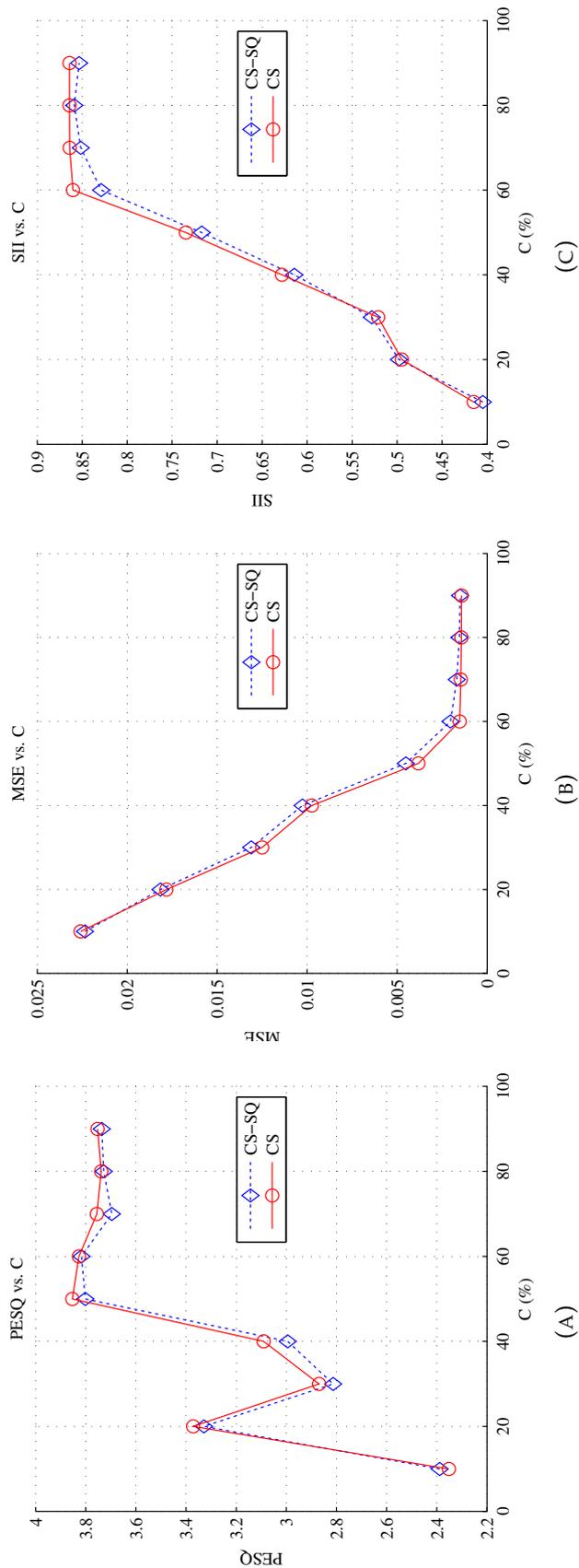


FIGURE 3.12: The effect of scalar quantization on compressed sensing recovery. Note that the performances of speech codec based on compressed sensing only (without quantization) and those of the speech codec based on scalar quantization of CS measurements (with $\text{nbps} = 4$) are similar for various compression ratios C .

at 1.6 kbit/s, and we varied C from 10% to 90%. Here, the number of bits per vector ($nbpv$) necessary to achieve 1.6 kbit/s was computed for every value of C for creating the codebooks. The performance scores of VQ, SVQ, MSVQ, and S-MSVQ are shown in Figs. 3.13, 3.14, 3.15, and 3.16, respectively. In the second case, the compression ratio was fixed to 10%, and the number of bits per vector ($nbpv$) was varied from 10 to 18 bits/vector. The results are shown in Figs. 3.17, 3.18, 3.19, and 3.20. A comparison between the different CS-based speech codecs (using the different quantization methods) was also performed. Note that we have limited the tests to these compression ratios and BRs due to memory constraint as the complexity of LBG algorithm increases exponentially with the number of bits per vector $nbpv$.

Case 1: $BR = 1.6 \text{ kbit/s}$ and $10\% \leq C \leq 90\%$

In this case, we study the effect of the compression ratio (C) on the quality and intelligibility of the recovered speech after compression using different quantization methods with CS measurements.

Performance using VQ Fig. 3.13 presents the average PESQ, MSE, and SII scores of the recovered speech after compression using VQ of CS measurements in the case of $BR=1.6 \text{ kbit/s}$ when different values of C are considered. *Fair* and *good* PESQ scores ($2.3002 \leq PESQ \leq 2.6897$) are achieved with the studied C (Fig. 3.13(A)). For example, for $C = 20\%$ the $PESQ = 2.58$, and for $C = 70\%$ the $PESQ = 2.69$. The highest PESQ score ($PESQ=2.6897$) is achieved when $C = 70\%$.

From Fig. 3.13(B), one can notice that low MSE scores are achieved. They are close to 0.02 for all the studied values of C .

Poor average intelligibility scores are achieved, as depicted in Fig. 3.13(C), which shows low average SII values ($0.3733 \leq SII \leq 0.4231$). The highest average SII score is achieved for $C = 40\%$.

Performance using SVQ The compressed sensing-based speech codec using SVQ was evaluated for two and three splits SVQ. Results in Fig. 3.14 demonstrate that this method performs competitively for two-splits and three-splits SVQ. *Fair* and *good* PESQ scores ($2.1547 \leq PESQ \leq 2.7123$ for two-splits SVQ and $2.0683 \leq PESQ \leq 2.8416$ for three-splits SVQ) are achieved with the studied C as observed in Fig. 3.14(A). For example, for $C = 30\%$ the $PESQ = 2.5344$ using two-splits SVQ and $PESQ= 2.6179$ using three-splits SVQ. For $C = 70\%$ the $PESQ = 2.6261$ using two-splits SVQ and $PESQ= 2.5455$ using three-splits SVQ. The best average PESQ score (2.8416) is achieved using three-splits SVQ for $C = 50\%$.

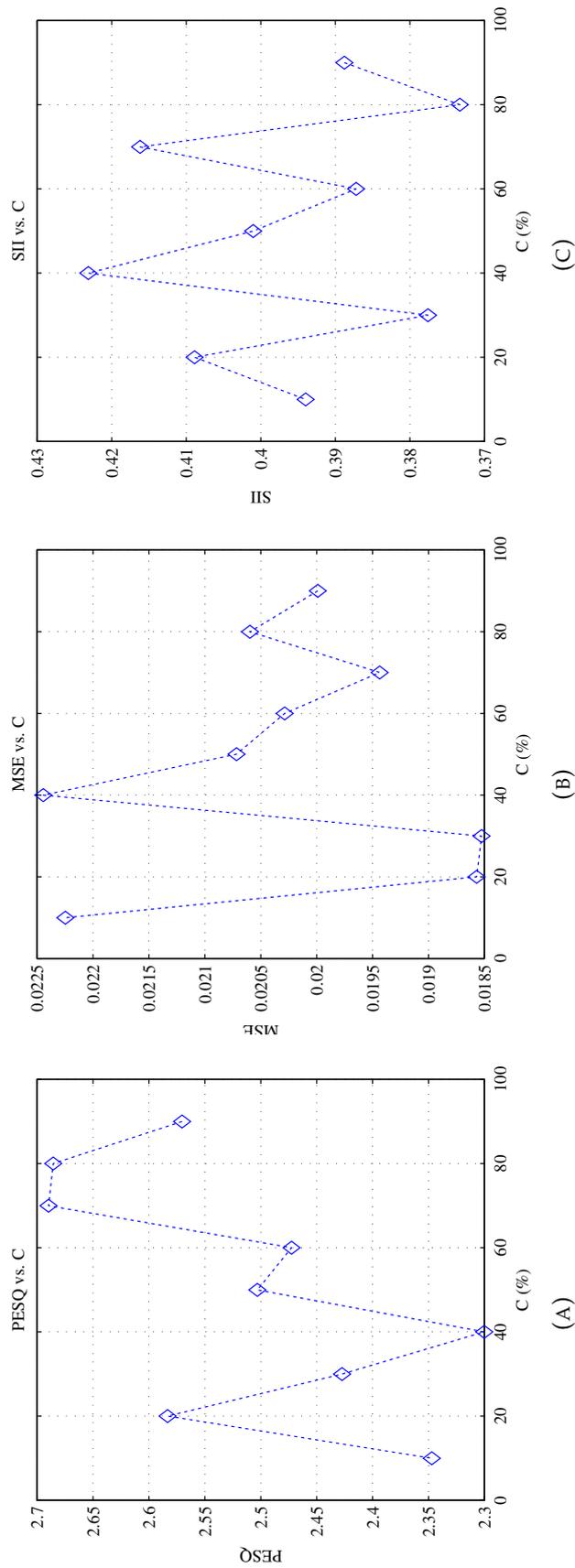


FIGURE 3.13: Performance of speech compression through vector quantization of compressed sensing measurements. Note that the bitrate is fixed to 1.6 kbit/s the compression ratio (C) is varying from 10% to 90%.

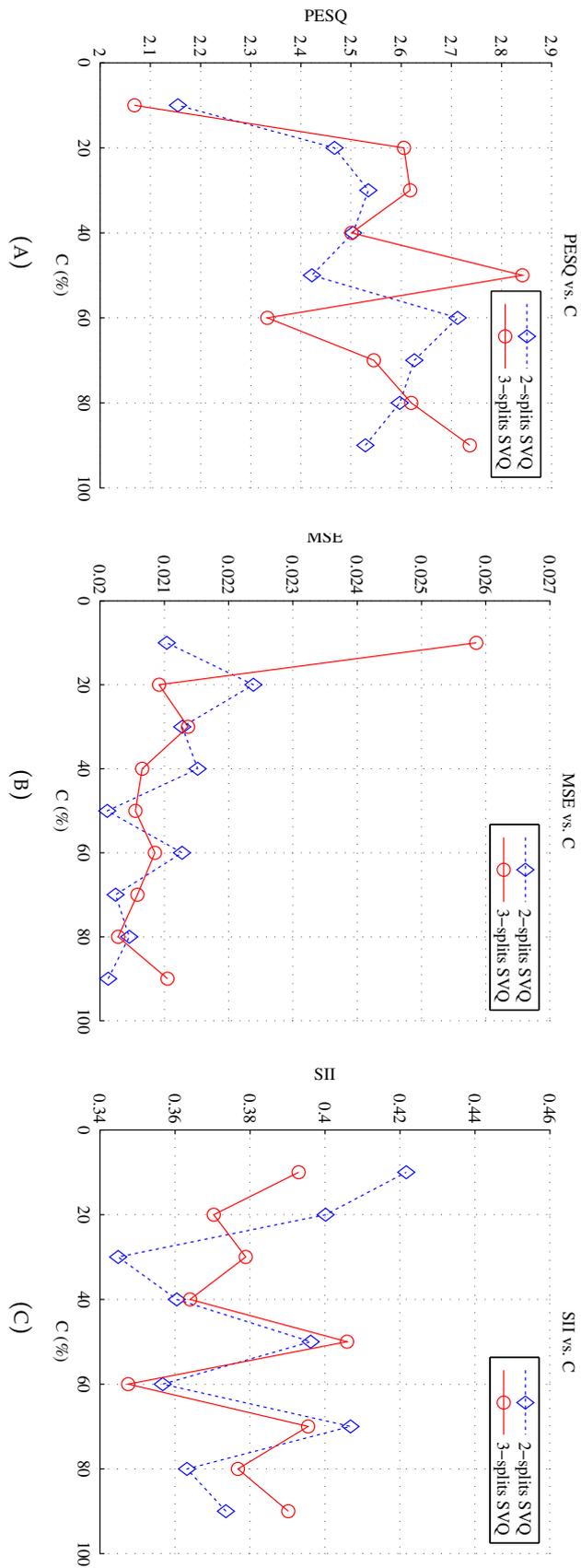


FIGURE 3.14: Performance of speech compression through split vector quantization of compressed sensing measurements. Note that the bitrate is fixed to 1.6 kbit/s the compression ratio (C) is varying from 10% to 90%.

From Fig. 3.14(B), one can notice that low MSE scores are achieved. They are close to 0.02 for all the studied values of C .

Poor average intelligibility scores are obtained (Fig. 3.14(C)). From this figure, one can note that the two codecs achieve low average SII values ($0.3448 \leq \text{SII} \leq 0.4216$ for two-splits SVQ and $0.3475 \leq \text{SII} \leq 0.4059$ for three-splits SVQ). The highest average SII score (SII=0.4216) is achieved using two-splits SVQ for $C = 10\%$.

Performance using MSVQ Two and three-stages MSVQ of CS measurements were performed for speech coding. As depicted in Fig. 3.15, performances in the case of two-stages MSVQ are close to those of three-stages MSVQ for various values of C . One can notice from Fig. 3.15(A) that *fair* and *good* PESQ scores ($2.2698 \leq \text{PESQ} \leq 2.8560$ for two-stages MSVQ and $2.2999 \leq \text{PESQ} \leq 2.7443$ for three-stages MSVQ) are achieved with the studied C . For example, for $C = 30\%$ the PESQ = 2.5185 using two-stages MSVQ and PESQ= 2.5164 using three-stages MSVQ. For $C = 80\%$ the PESQ = 2.6831 using both two-splits MSVQ and three-stages MSVQ. The best average PESQ score (2.8560) is achieved using two-stages MSVQ for $C = 60\%$.

It is seen in Fig. 3.15(B) that low MSE scores, close to 0.02, are achieved for all the studied range of C .

As shown in Fig. 3.15(C), *poor* average intelligibility scores are achieved, which is interpreted by the low average SII values ($0.3706 \leq \text{SII} \leq 0.4300$ for two-stages MSVQ and $0.3553 \leq \text{SII} \leq 0.4113$ for three-stages MSVQ). The highest average SII score (SII=0.4300) is achieved using two-stages MSVQ for $C = 60\%$.

Performance using S-MSVQ Speech coding using S-MSVQ of CS measurements was done for various numbers of splits and stages. One can notice from Fig. 3.16 that this speech codec performs competitively for the different numbers of splits/stages, with the studied values of C . It is noted in Fig. 3.16(A) that *fair* and *good* PESQ scores are achieved. For instance, for $C = 50\%$ the PESQ = 2.6636 using two-stages two-splits S-MSVQ, PESQ = 2.5280 using two-stages three-splits S-MSVQ, PESQ = 2.7542 using three-stages two-splits S-MSVQ, and PESQ = 2.6664 using three-stages three-splits S-MSVQ. The best average PESQ score (2.8580) is achieved using two-stages three-splits S-MSVQ for $C = 20\%$.

From Fig. 3.16(B), it is observed that low MSE scores, close to 0.02, are achieved for all the studied range of C .

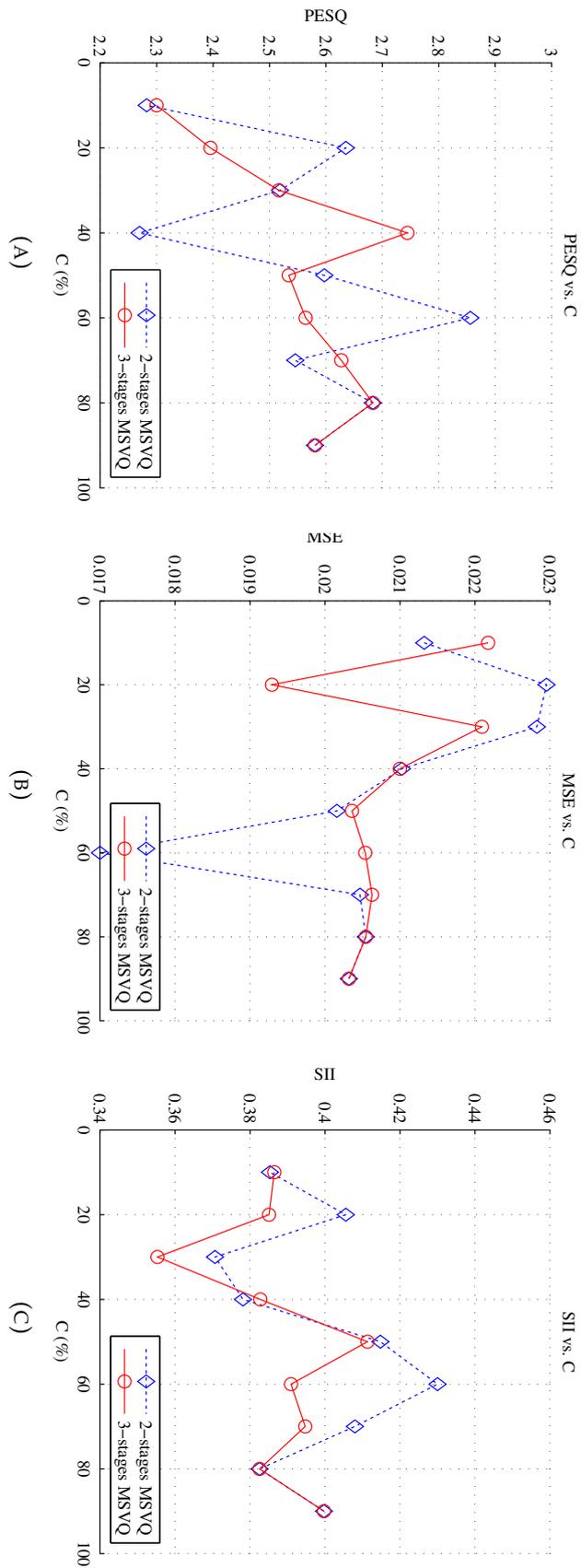


FIGURE 3.15: Performance of speech compression through multistage vector quantization of compressed sensing measurements. Note that the bitrate is fixed to 1.6 kbit/s the compression ratio (C) is varying from 10% to 90%.

Poor average intelligibility scores are achieved, which is interpreted by the low SII values observed in Fig. 3.16(C). The highest average SII score (SII=0.4320) is achieved using two-stages three-splits S-MSVQ for $C = 20\%$.

Comparison of the different CS-based codecs Table 3.1 compares the average PESQ, MSE, and SII scores of the recovered speech after compression using SQ, VQ, three-splits SVQ, two-stages MSVQ, and two-stages three-splits S-MSVQ of CS measurements. The execution times of quantization/dequantization are also listed in this table. Note that for SQ, we provide only the results for $C = 10\%$ as the minimum bitrate achieved (1.6 kbit/s) for the studied range of C is obtained when $C = 10\%$ and $nbps = 1$ bit. Higher values of C lead to higher BRs. As highlighted in this table, the proposed speech codecs (CS with SVQ/S-MSVQ) achieve *good* average PESQ scores for low bitrate ($BR = 1.6$ kbit/s). They are better than the values obtained for SQ and VQ. To be specific, the highest PESQ scores for SVQ and S-MSVQ (values in bold in the table); namely 2.842 and 2.858, respectively (obtained for $C = 50\%$ using SVQ and $C = 20\%$ using S-MSVQ) are better than the highest PESQ score of VQ which is 2.690 (obtained for $C = 70\%$) and the PESQ score of SQ which is 2.273. Moreover, S-MSVQ leads to the best average PESQ score over all the other quantization methods including MSVQ.

For the MSE performance, one can note that the values obtained using vector quantizations (VQ, SVQ, MSVQ, S-MSVQ) are better than those obtained using SQ. Despite S-MSVQ and MSVQ perform better than the other quantizations, the MSE performance of all of them is comparable, i.e., there is not a big difference.

Poor intelligibility scores are obtained for all the studied quantizations. Vector quantizations (VQ, SVQ, MSVQ, and S-MSVQ) perform better than SQ. The best performance is achieved using S-MSVQ. By way of illustration, the best average intelligibility score achieved using S-MSVQ, which is 0.432 (obtained for $C = 20\%$), is better than the values obtained using SQ (0.384), VQ (0.423), SVQ (0.406), and MSVQ (0.430).

In addition to quality and intelligibility assessment, we compared the execution times. We notice (from the same table) that the execution time of SQ is the lowest, this can be justified by the fact that SQ has less complexity. We also notice that the execution times of SVQ, MSVQ, and S-MSVQ (values in bold) are lower than that of VQ. The lowest execution time, over vector quantizations, is achieved using S-MSVQ when $C = 80\%$. This is because S-MSVQ has less complexity compared to the other vector quantizations (VQ, SVQ, and S-MSVQ) as it shares the bits available ($nbps$) between the splits and stages.

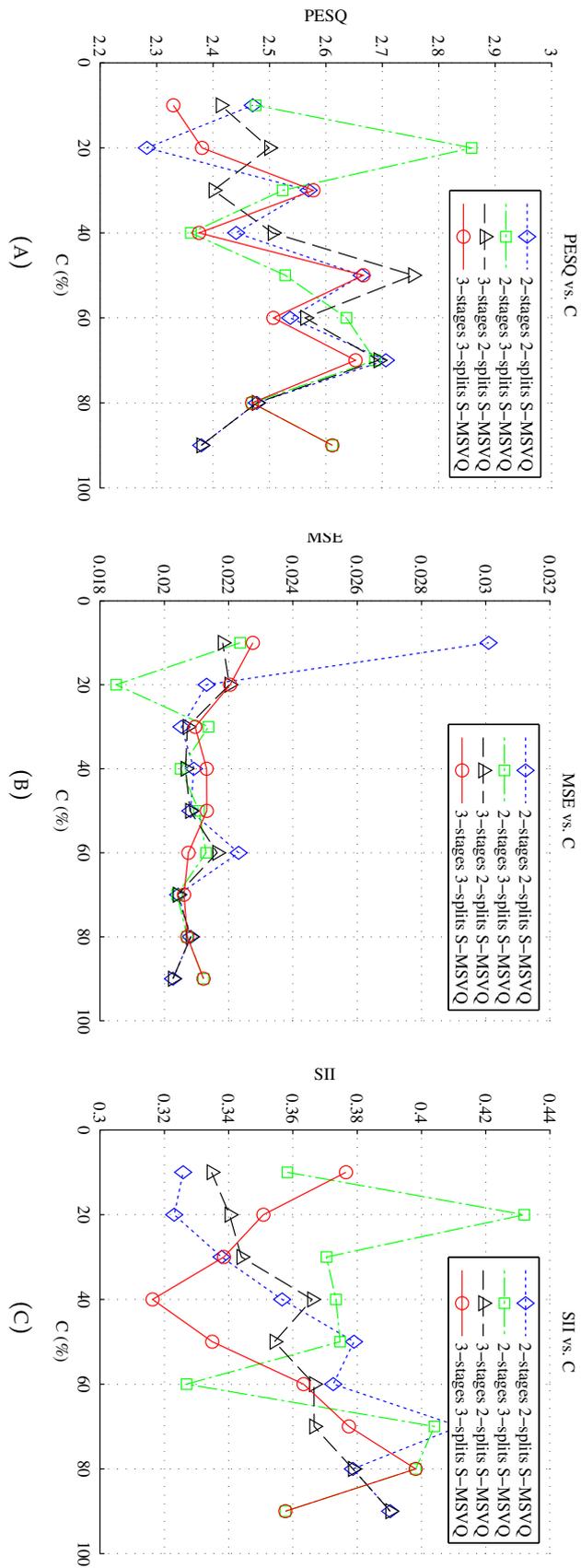


FIGURE 3.16: Performance of speech compression through split-multistage vector quantization of compressed sensing measurements. Note that the bitrate is fixed to 1.6 kbit/s the compression ratio (C) is varying from 10% to 90%.

TABLE 3.1: The average performance scores of the recovered speech signals after compression ($BR = 1.6$ kbit/s) using different quantization techniques (namely, scalar quantization, vector quantization, three-splits split vector quantization, two-stages multistage vector quantization, and two-stages three-splits split-multistage vector quantization) with CS measurements. *Note that the bitrate is fixed to 1.6 kbit/s the compression ratio (C) is varying from 10% to 90%. The best scores for each quantization method are shown in bold.*

Measure	C (%)	Quantization Method				
		SQ	VQ	SVQ	MSVQ	S-MSVQ
PESQ	10	2.273	2.347	2.068	2.282	2.475
	20	—	2.583	2.606	2.635	2.858
	30	—	2.427	2.618	2.519	2.523
	40	—	2.300	2.501	2.270	2.361
	50	—	2.503	2.842	2.297	2.528
	60	—	2.472	2.333	2.856	2.636
	70	—	2.690	2.546	2.546	2.687
	80	—	2.686	2.620	2.683	2.469
	90	—	2.570	2.737	2.581	2.611
MSE	10	0.022	0.022	0.026	0.021	0.022
	20	—	0.019	0.021	0.023	0.018
	30	—	0.019	0.021	0.023	0.021
	40	—	0.022	0.021	0.021	0.021
	50	—	0.021	0.021	0.020	0.021
	60	—	0.020	0.021	0.017	0.021
	70	—	0.019	0.021	0.021	0.020
	80	—	0.021	0.020	0.021	0.021
	90	—	0.020	0.021	0.020	0.021
SII	10	0.384	0.394	0.393	0.385	0.358
	20	—	0.409	0.370	0.406	0.432
	30	—	0.378	0.379	0.371	0.370
	40	—	0.423	0.364	0.378	0.373
	50	—	0.401	0.406	0.415	0.375
	60	—	0.387	0.348	0.430	0.327
	70	—	0.416	0.396	0.408	0.404
	80	—	0.373	0.377	0.383	0.398
	90	—	0.388	0.390	0.400	0.358
Time (s)	10	0.0001	0.054	0.012	0.009	0.013
	20	—	0.027	0.012	0.010	0.012
	30	—	0.026	0.011	0.014	0.008
	40	—	0.027	0.011	0.011	0.009
	50	—	0.029	0.013	0.008	0.012
	60	—	0.028	0.008	0.012	0.010
	70	—	0.026	0.009	0.009	0.009
	80	—	0.030	0.012	0.009	0.007
	90	—	0.026	0.012	0.011	0.011

Case 2: $C = 10\%$ and $10 \leq nbpv \leq 90\%$ bits/vector

In the second case, we focus on the effect of the number of bits per vector ($nbpv$) on the quality and intelligibility of the recovered speech after compression using different quantization methods with CS measurements. Note that varying the number of bits per vector ($nbpv$) influences the bitrate, as the latter is computed, for vector quantizations, as follows

$$BR = nbpv \times N_v / T_f \quad (3.6)$$

where, $nbpv$ is the number of bits per vector, T_f is the speech frame duration, and N_v is the number of vectors in the frame.

$$N_v = M / dim \quad (3.7)$$

where M is the CS measurements number, and dim is the number of samples (CS measurements) included in a vector considered for vector quantization. From Eqs. (3.4), (3.6), and (3.7), the bitrate can be computed as

$$BR = C \times f_s \times nbpv / dim \quad (3.8)$$

where f_s is the sampling frequency of the speech. It is also noticed from the latter equation (compared to Eq. (3.5)) that the lowest bitrate than can be achieved by VQ (considering $nbpv = 1$) is dim times lesser than the lowest bitrate that can be achieved using SQ (considering $nbps = 1$). Returning to our experiment, where $dim = 16$, $f_s = 16$ kHz, $C = 10\%$. Considering $10 \leq nbpv \leq 18$ bits/vector means that $1 \leq BR \leq 1.8$ kbit/s. This is theoretically. In practice M may not to be multiple of dim ; hence N_v can be chosen as the nearest integer greater than or equal to M/dim . Consequently, slightly higher bitrates than the theoretic ones are achieved. For our tests $1.2 \leq BR \leq 2.16$ kbit/s.

Performance using VQ Fig. 3.17 presents the average PESQ, MSE, and SII scores of the recovered speech after compression using VQ of CS measurements in the case of $C = 10\%$ when different values of $nbpv$ are considered. Fair PESQ scores ($2.3357 \leq PESQ \leq 2.3869$) are achieved with the studied $nbpv$ (Fig. 3.17(A)). For example, for $nbpv = 14$ bits/vector, the PESQ = 2.3472. The highest average PESQ score (PESQ=2.3869) is achieved when $nbpv = 12$ bits/vector.

We notice from Fig. 3.17(B) that low MSE scores are achieved. They are close to 0.02 for all the studied values of $nbpv$.

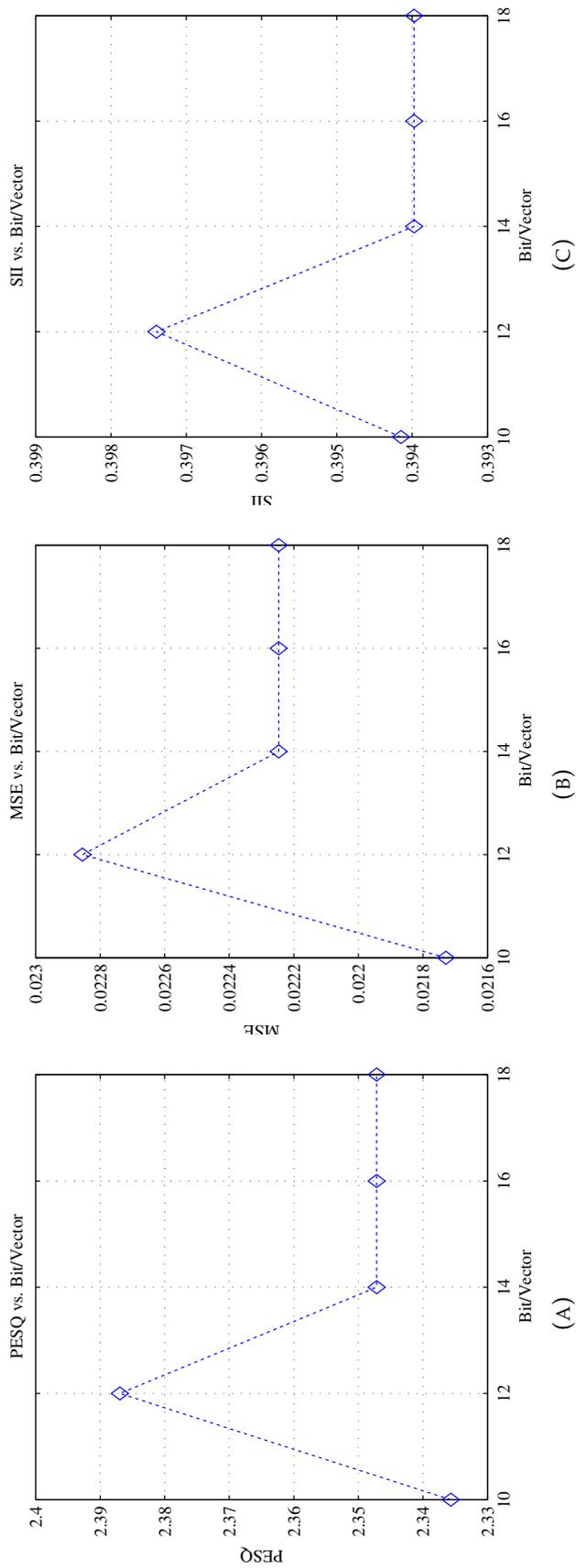


FIGURE 3.17: Performance of speech compression through vector quantization of compressed sensing measurements. Note that the compression ratio (C) is fixed to 10% and the number of bits per vector is varying from 10 to 18 bits/vector.

Poor average intelligibility scores are achieved, as depicted in Fig. 3.17(C), which shows low average SII values ($0.3940 \leq \text{SII} \leq 0.3974$). The highest average SII score is achieved for $nbpv = 12$ bits/vector. Note that there is not a big difference between the performance results of the studied $nbpv$ as the difference between the chosen numbers of bits per vector is small (only two bits). This remark is kept for the subsequent results.

Performance using SVQ The proposed CS-based speech codec (using SVQ) was evaluated for two and three splits SVQ. Results in Fig. 3.18 demonstrate that this method performs competitively for two-splits and three-splits SVQ with regard to the obtained PESQ scores. Fair and good PESQ scores ($2.1309 \leq \text{PESQ} \leq 2.4696$ for two-splits SVQ and $2.0683 \leq \text{PESQ} \leq 2.7629$ for three-splits SVQ) are achieved with the studied $nbpv$ as observed in Fig. 3.18(A). For example, for $nbpv = 10$ bits/vector, the PESQ = 2.4696 using two-splits SVQ and PESQ= 2.7629 using three-splits SVQ. For $nbpv = 12$ bits/vector, the PESQ = 2.4075 using two-splits SVQ and PESQ= 2.2245 using three-splits SVQ. The best average PESQ score (2.7629) is achieved using three-splits SVQ for $nbpv = 10$ bits/vector (i.e., $BR = 1.2$ kbit/s).

From Fig. 3.18(B), one can notice that low MSE scores are achieved. They are close to 0.02 for all the studied values of $nbpv$. The MSE values obtained using two-splits SVQ are lower than those obtained using three-splits SVQ.

From Fig. 3.18(C), we observe that poor average intelligibility scores are obtained. The two codecs achieve low average SII values, to clarify $0.3366 \leq \text{SII} \leq 0.4407$ for two-splits SVQ, and $0.2866 \leq \text{SII} \leq 0.3930$ for three-splits SVQ. The SII values obtained using two-splits SVQ are higher than those obtained using three-splits SVQ for all the studied range of $nbpv$. The highest average SII score (SII=0.4407) is achieved using two-splits SVQ for $nbpv = 18$ bits/vector.

Performance using MSVQ Fig. 3.19 reports the performance of a CS-based speech codec using MSVQ, where two and three stages are considered. As can be seen in this figure, this codec performs competitively for two-stages and three-stages MSVQ. From Fig. 3.19(A), we can note that fair and good PESQ scores are achieved with the studied $nbpv$. Indeed, $2.2573 \leq \text{PESQ} \leq 2.7910$ for two-stages MSVQ, and $2.2118 \leq \text{PESQ} \leq 2.5287$ for three-stages MSVQ. To explain further, for $nbpv = 12$ bits/vector, the PESQ = 2.3879 using two-stages MSVQ and PESQ= 2.5287 using three-stages MSVQ. For $nbpv = 14$ bits/vector, the PESQ = 2.7910 using two-stages MSVQ and PESQ= 2.2999 using three-stages MSVQ. The best average PESQ score (2.7910) is achieved using two-stages MSVQ for $nbpv = 14$ bits/vector (i.e., $BR = 1.68$ kbit/s).

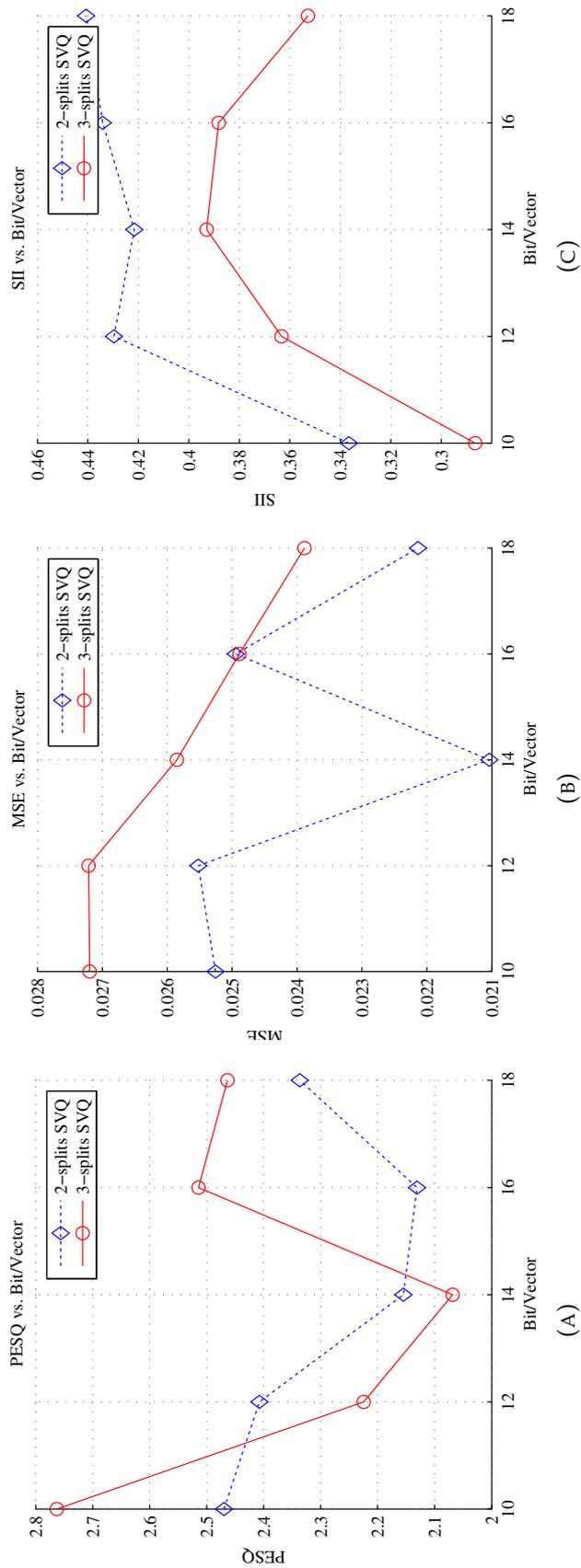


FIGURE 3.18: Performance of speech compression through split vector quantization of compressed sensing measurements. Note that the compression ratio (C) is fixed to 10%, and the number of bits per vector is varying from 10 to 18 bits/vector, which is equivalent to $1.2 \leq BR \leq 2.16$ kbit/s.

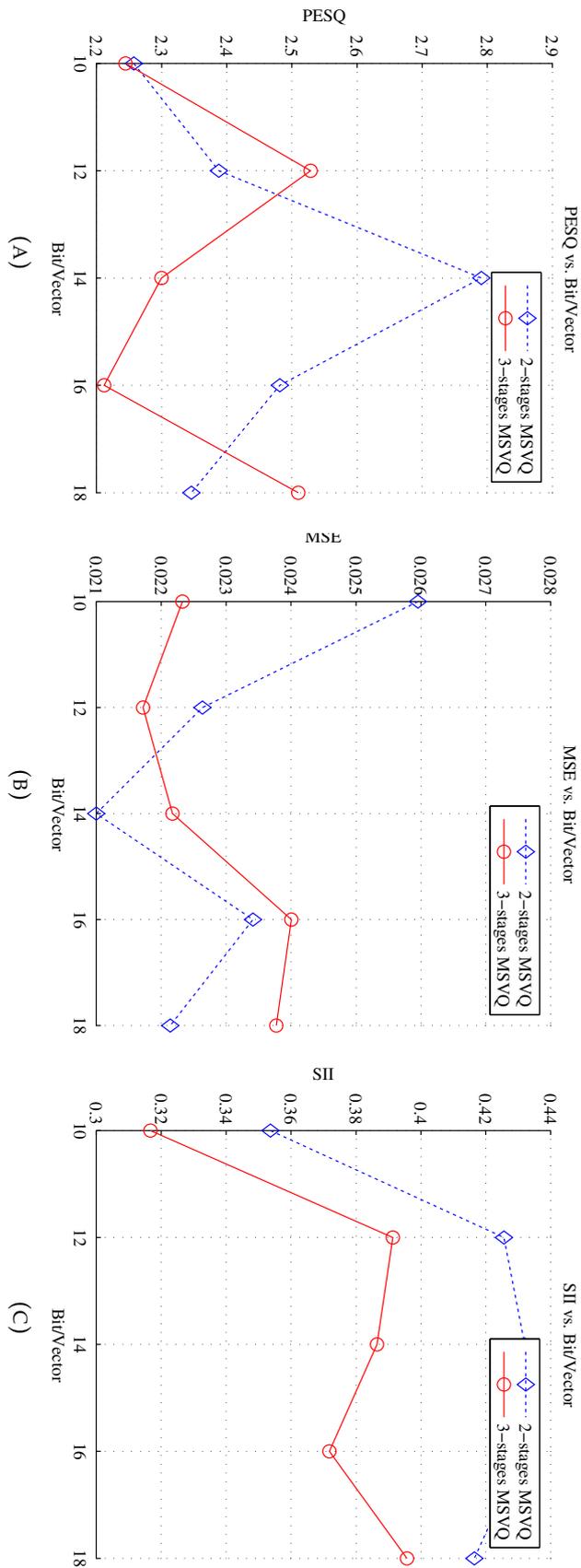


FIGURE 3.19: Performance of speech compression through multistage vector quantization of compressed sensing measurements. Note that the compression ratio (C) is fixed to 10% and the number of bits per vector is varying from 10 to 18 bits/vector.

Low MSE scores, close to 0.02, are achieved for all the studied values of $nbpv$ as depicted in Fig. 3.19(B). The average MSE values obtained using three-stages MSVQ are lower than those obtained using two-stages MSVQ for $nbpv \leq 13$ bits/vector. For $nbpv > 13$ the CS-speech codec using two-stages MSVQ achieves low MSE values.

As illustrated in Fig. 3.19(C), *poor* average intelligibility scores are obtained. To clarify, $0.3537 \leq \text{SII} \leq 0.4320$ for two-stages MSVQ, and $0.3167 \leq \text{SII} \leq 0.3957$ for three-stages MSVQ. The SII values obtained using two-stages MSVQ are higher than those obtained using three-stages MSVQ for all the studied range of $nbpv$. The highest average SII score (SII=0.4320) is achieved using two-stages MSVQ for $nbpv = 14$ bits/vector.

Performance using S-MSVQ The other proposed speech codec based on CS and S-MSVQ was evaluated for various numbers of splits and stages. The performance results highlighted in Fig. 3.20 reveal that this speech codec performs competitively for the different numbers of splits/stages, with the studied values of $nbpv$. We observe from Fig. 3.20(A) that *fair* PESQ scores are achieved for two-stages two-splits S-MSVQ, three-stages two-splits S-MSVQ, and three-stages three-splits S-MSVQ, over the studied range of $nbpv$. When two-stages three-splits S-MSVQ is used, *fair* and *good* average PESQ scores are obtained. For example, for $nbpv = 12$ bits/vector, the PESQ = 2.3022 using two-stages two-splits S-MSVQ, PESQ = 2.8030 using two-stages three-splits S-MSVQ, PESQ = 2.2669 using three-stages two-splits S-MSVQ, and PESQ = 2.4606 using three-stages three-splits S-MSVQ. The best average PESQ score (2.8030) is achieved using two-stages three-splits S-MSVQ for $nbpv = 12$ bits/vector. It is also apparent from this figure that for almost all the values of $nbpv$, two-stages three-splits S-MSVQ gives the highest average PESQ scores.

As shown in Fig. 3.20(B), low MSE scores, close to 0.02, are achieved for all the studied range of $nbpv$ when two-stages three-splits S-MSVQ, three-stages two-splits S-MSVQ, or three-stages three-splits S-MSVQ are used.

Poor average intelligibility scores are achieved, which is interpreted by the low SII values observed in Fig. 3.20(C). The highest average SII score (SII=0.4340) is achieved using two-stages three-splits S-MSVQ for $nbpv = 12$ bits/vector.

Comparison of the different CS-based codecs Table 3.2 compares the average PESQ, MSE, and SII scores of the recovered speech after compression using SQ, VQ, three-splits SVQ, two-stages MSVQ, and two-stages three-splits S-MSVQ of CS measurements, in the case of $C = 10\%$ and $10 \leq nbpv \leq 18$ bits/vector. The execution times of quantization/dequantization are also reported. For SQ, only the

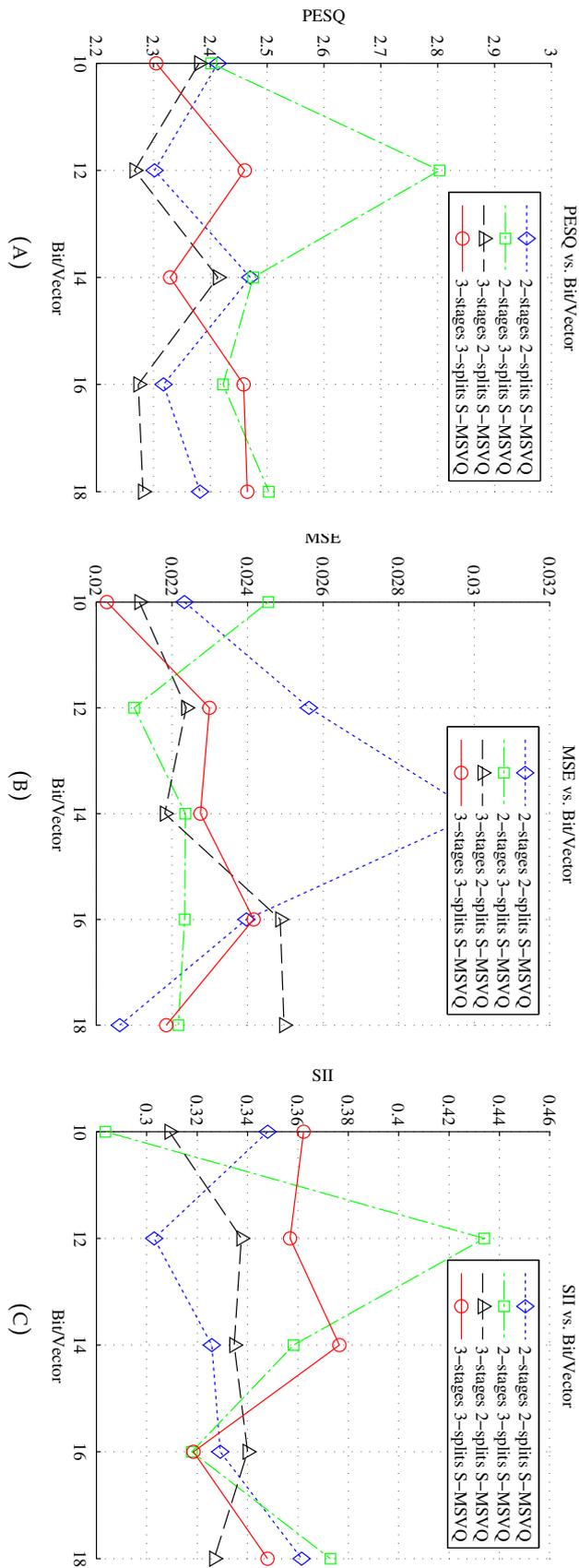


FIGURE 3.20: Performance of speech compression through split-multistage vector quantization of compressed sensing measurements. Note that the compression ratio (C) is fixed to 10% and the number of bits per vector is varying from 10 to 18 bits/vector.

results for theoretic bitrate $BR = 1.6$ kbit/s (corresponding to $nbpv = 16$ bits/vector) are provided as the minimum bitrate achieved (1.6 kbit/s) using SQ for the studied ranges of $nbps$ and C is obtained when $C = 10\%$ and $nbps = 1$ bit. Higher values of $nbps$ lead to higher BRs .

TABLE 3.2: The average performance scores of the recovered speech signals after compression using different quantization techniques (namely, scalar quantization, vector quantization, three-splits vector quantization, two-stages vector quantization, and two-stages three-splits vector quantization) with compressed sensing measurements. Note that the compression ratio (C) is fixed to 10% and the number of bits per vector are varying from 10 to 18 bits.

Measure	Bit/Vector	Quantization Method				
		SQ	VQ	SVQ	MSVQ	S-MSVQ
PESQ	10	–	2.336	2.763	2.257	2.402
	12	–	2.387	2.225	2.388	2.803
	14	–	2.347	2.068	2.791	2.475
	16	2.273	2.347	2.514	2.481	2.422
	18	–	2.347	2.463	2.346	2.503
MSE	10	–	0.022	0.027	0.026	0.025
	12	–	0.023	0.027	0.023	0.021
	14	–	0.022	0.026	0.021	0.022
	16	0.022	0.022	0.025	0.023	0.022
	18	–	0.022	0.024	0.022	0.022
SII	10	–	0.394	0.287	0.354	0.284
	12	–	0.397	0.363	0.426	0.434
	14	–	0.394	0.393	0.432	0.358
	16	0.384	0.394	0.388	0.431	0.318
	18	–	0.394	0.353	0.417	0.373
Time (s)	10	–	0.012	0.009	0.009	0.006
	12	–	0.010	0.008	0.008	0.007
	14	–	0.011	0.007	0.007	0.008
	16	0.0001	0.018	0.007	0.008	0.007
	18	–	0.018	0.008	0.014	0.009

Results in this table prove that, the proposed speech codecs (CS with SVQ/S-MSVQ) achieves *good* average PESQ scores for low compression ratio ($C = 10\%$). They are better than the values obtained for SQ and VQ. To explain further, the highest PESQ scores for SVQ and S-MSVQ (values in bold in the table); namely 2.763 and 2.803, respectively (obtained for $nbpv = 10$ bits/vector using SVQ and $nbpv = 12$ bits/vector using S-MSVQ) are better than the highest PESQ score achieved using VQ which is 2.387 (obtained for $nbpv = 12$ bits/vector) and the PESQ score achieved

using SQ which is 2.273. One can also notice that using S-MSVQ to quantize CS measurements results in the best average PESQ score over all the other quantization methods including MSVQ.

Regarding the MSE performance, we notice that S-MSVQ and MSVQ perform better than the other quantizations (SQ, VQ, SVQ). However, the MSE performance of all of them is generally comparable.

As for the intelligibility performance, we find that *poor* intelligibility scores are obtained for all the studied quantizations. One can note from the results reported in the table that vector quantizations (VQ, SVQ, MSVQ, and S-MSVQ) perform better than SQ. In addition, we notice that the intelligibility scores achieved using MSVQ and S-MSVQ are better than those obtained using VQ and SVQ. It is also apparent that the best performance is achieved using S-MSVQ. To be specific, the best average intelligibility score achieved using S-MSVQ, which is 0.434 (obtained for $nbpv = 12$ bits/vector), is better than the values obtained using SQ (0.384), VQ (0.397), SVQ (0.393), and MSVQ (0.432).

For the execution times comparison, one can note (from the same table) that the execution time of SQ is the lowest. It is also observed that the execution times of SVQ, MSVQ, and S-MSVQ (values in bold) are lower than that of VQ. The lowest execution time, over vector quantizations, is achieved using S-MSVQ when $nbpv = 10$ bits/vector. This is because S-MSVQ has less complexity compared to the other vector quantizations as stated before.

Comparison with Other Compressed Sensing-Based Codecs

In addition to the comparison with CS-based speech coding using SQ, VQ, and MSVQ of CS measurements, we compared the performance of our speech codec with other CS-based speech coding; namely the works of (Giacobello et al., 2012) and (Al-Azawi and Gaze, 2018). These two works have not used quantization for CS measurements. In (Giacobello et al., 2012), the authors have introduced sparsity in a linear prediction framework to improve the linear prediction coding performances. (Al-Azawi and Gaze, 2018) has introduced a combined compression/encryption method for speech using CS. Contourlet transform increases the sparsity whereas the chaotic system generates the sensing matrix in the proposed approach. The average PESQ scores shown in Table 3.3 demonstrate that our proposed methods perform better than the method proposed in (Al-Azawi and Gaze, 2018). In the latter work, the authors didn't use a quantization scheme. They obtained a PESQ score of 2.73 for a $C = 40\%$. Although we used a quantization scheme (which engenders a quantization

error) and lower $C = 10\%$, our PESQ values (2.76 for SVQ and 2.80 for S-MSVQ) are better. The PESQ value obtained by the method proposed in (Giacobello et al., 2012), for a $BR = 10.1$ kbits/s and $C = 50\%$, is close to 4.13 ± 0.13 . Despite this value is higher than our PESQ scores, our methods lead to good results for significantly lower bitrates (1.20 and 1.44 kbit/s) and lower compression ratio ($C = 10\%$). Our methods can achieve higher scores for higher bitrates.

TABLE 3.3: Performance measures of the two proposed speech codecs (compressed sensing with three-splits split vector quantization and compressed sensing with two-stages three-splits split-multistage vector quantization)

Method	Compression Ratio	PESQ
Proposed 1 (CS with SVQ)	10% ($BR = 1.20$ kbit/s)	2.76
Proposed 2 (CS with S-MSVQ)	10% ($BR = 1.44$ kbit/s)	2.80
(Giacobello et al., 2012)	50% ($BR = 10.1$ kbit/s)	4.13 ± 0.13
(Al-Azawi and Gaze, 2018)	40%	2.73

Comparison with currently used compression standards

Our CS-codec using S-MSVQ was also compared with two currently used speech coding standards, namely AMR and AMR-WB (Cheraitia and Bouzid, 2014). The 2nd proposed CS-codec was tested in noise free and in presence of additive Gaussian noise (AWGN) noise. Table 3.4 presents the PESQ performance of the recovered speech from our CS-codec (using S-MSVQ), AMR, and AMR-WB in noise free and AWGN environment. The proposed CS-speech coding allows acceptable speech

TABLE 3.4: PESQ comparison of the proposed CS-codec with AMR and AMR-WB in noise free and additive white Gaussian noise environment.

Method	BR (kbit/s)	Noise free	20 dB	10 dB	5 dB
Proposed 2 (CS with S-MSVQ)	8.85	3.14	3.02	2.95	2.72
AMR (Giacobello et al., 2012)	10.2	4.02 ± 0.11	3.25 ± 0.19	2.76 ± 0.23	–
AMR-WB (Cheraitia and Bouzid, 2014)	12.65	3.77	–	3.18	1.35

quality compared to AMR and AMR-WB. To be more specific, in the noisy case, the average PESQ scores obtained using the proposed CS-codec are close to those of AMR even with smaller BR . Additionally, in 5 dB white Gaussian noise, the average PESQ score obtained using the proposed CS-codec for a $BR = 8.85$ kbit/s (PESQ= 2.72) is better than that obtained using AMR-WB at $BR = 12.65$ kbit/s (PESQ= 1.35). One can also notice that our PESQ scores decrease at the lowest rate, which means that

our CS-codec is robust to the SNR fall. The latter is achieved thanks to the robustness of ℓ_1 minimization, and the use of wavelet thresholding (performed to enhance the sparsity) which is considered as a denoising method in many applications, e.g., (Sekkate, Khalil, and Adib, 2019).

3.8 Conclusion

As this thesis falls in the area of speech compression and communication, we have presented in this chapter two propositions of CS-based speech codec using two different quantization methods; namely, SVQ and S-MSVQ with CS measurements. Additionally, we have showed results for speech when SQ, VQ, and MSVQ are used.

Our results casts light on the findings summarized below:

- Using SQ to quantize CS measurements for speech compression allows *good* PESQ scores using only 1 bit/sample with $C = 20\%$, and *fair* intelligibility scores using only 2 bits/sample and $C = 20\%$. Considering 4 bits/sample, SQ does not affect the quality of CS reconstruction. Moreover, choosing $nbps = 4$ bits/sample and $C = 20\%$, a trade-off between the bitrate and the speech quality is achieved. However, better results for lower bitrates can be achieved using vector quantizations. Results provide also evidence that SQ permits the lowest execution time.
- In the case of vector quantizations, for a fixed bitrate ($BR = 1.6$ kbit/s in our case), the rising of C does not mean that better quality/intelligibility scores are achieved. The value of C that leads to the highest scores can be selected empirically.
- Likewise, for a fixed compression ratio ($C = 10\%$ in our case), the rising of the number of bits per vector does not mean that better quality/intelligibility scores are achieved. The value of $nbpv$ that leads to the best performance can be selected empirically. This is very likely due to the fact that the difference between the studied values of $nbpv$ is small.
- Yet, the rising of the number of splits and/or stages in the case of SVQ, MSVQ, and S-MSVQ does not mean that better performance is achieved. Each method performs competitively with regard of the number of splits and/or stages. One can empirically choose the number of splits and/or stages that leads to a compromise between quality and complexity.

-
- For the studied bitrate ($BR = 1.6$ kbit/s), all the studied methods (including SQ) give *poor* intelligibility scores. Better values can be achieved for higher bitrates.
 - Moreover, using our proposed speech codecs (CS with SVQ/S-MSVQ), *good* quality performance for lower bitrate is achieved with an acceptable execution time compared to the CS-speech codecs using SQ.
 - The best quality performance is achieved using S-MSVQ. The latter uses less execution time than the other vector quantizations.

In the following chapter, a novel mobile communication system exploiting CS is developed.

Chapter 4

Designing a New Mobile Communication System by Exploiting Compressed Sensing

4.1 Introduction

The consumption of physical resources used for mobile communications depends on the data volume. The latter continues to increase; hence the need to improve the performance of systems used for transmission. In this perspective, we propose a new end-to-end communication system by using compressed sensing (CS) instead of the traditional Nyquist-Shannon sampling technique, for simultaneously acquiring and compressing speech (Haneche, Ouahabi, and Boudraa, 2019b). Our work considers all the transmission chain. In other words, it studies the behaviour of CS-source coding within the transmission chain when undergoing real communication-conditions. In addition, efficient techniques to mitigate the channel effects are chosen to meet the requirement of future communication systems such as 5G communications. In this chapter, the new CS-based mobile system is described, and results for speech communication are presented. Moreover, an alternative scheme is presented for lower bitrates (Haneche, Ouahabi, and Boudraa, 2019a).

4.2 Motivation

In current communication systems, signals are acquired and then source-coded (as shown in Fig. 4.1) to have a compressed form of speech in order to reduce network loading. High complexity standards, e.g., G.722.1, adaptive multi-rate wideband (AMR-WB, also normalized as ITU-T G.722.2), are used (Bessette et al., 2002; Cheraitia and Bouzid, 2014). These standards need several processes, which

increase the complexity of communication systems. Although these systems grant high bandwidth mobile service, it seems that they will not be able to meet the emerging demands of future mobile communications, where billions of devices require secured Gigabit wireless connectivity (Internet of Things). Researches on new technologies are in progress to encounter the current communication systems drawbacks summarized as follows (Wunder et al., 2015; Gao et al., 2018):

- Broadband services offered to mobile devices create inefficiencies in both spectrum usage and physical resources.
- The extensive growth of applications causes a very high energy consumption.
- Computational load is also an issue since a very big amount of data is acquired and then submitted to complex compression techniques.
- At high frequency ranges, robustness and reliability know an increasing risk in dramatic atmospheric conditions.
- As each person may be connected to many devices, an increasing risk of receiving an attack rises (security threats).

Future communication systems researches are not interested only by providing higher rates, but more interested by reducing the processing load and the memory requirements. Moreover, enhancing the battery life and the security are also a purpose. Considering the aforementioned shortcomings, we propose a low-cost solution based on sparse signal processing. The proposed design allows performing simultaneously the acquisition and the compression. We replaced the two blocks (in dashed line) in Fig. 4.1 by one CS-source coding module. The latter implicitly reduces the complexity of the communication system due to the use of simple quantization and binary encoding. In addition, it offers some level of security at zero-cost. This scheme is investigated in real communication conditions. The whole transmission chain is, hence, considered using efficient techniques to compensate the channel.

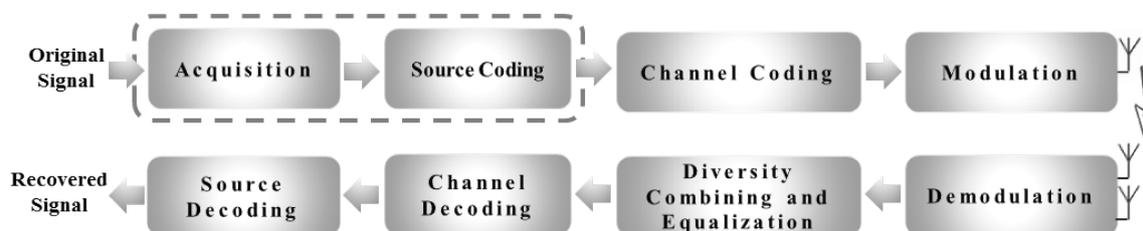


FIGURE 4.1: General diagram of communication systems

4.3 The Proposed End-to-End Communication System Model

Our goal is to contribute to the development of a mobile communication system by applying CS-source coding, performing multicarrier modulation, and using channel coding and receive antenna diversity to meet the requirement of future communication systems such as 5G communications. This entails mitigating the channel effects, optimizing spectral efficiency, and ensuring robustness and reliability while increasing the baud rate. The proposed system shows a simplified design, and allows reducing bit rates and processing load compared to actual communication systems based on, for example, G.722.1 and G.722.2 speech codecs. In addition, CS allows secure communications without additional costs.

The designed communication scheme is presented in Fig. 4.2 It includes, at the



FIGURE 4.2: General diagram of the proposed communication scheme (note that this system integrates compressed sensing-speech coding instead of the current codecs standards such as G.722.2)

transmission side, the following blocks:

- The proposed speech coder based on CS.
- Channel encoder using convolutional codes.
- Modulator using orthogonal frequency division multiplexing (OFDM).

More details about the proposed design are provided in the subsequent subsections.

4.3.1 Compressed Sensing-Based Source Coding

The new end-to-end communication system is proposed to increase transmission speed, robustness, and security in order to meet the requirements of mobile systems that know an exponentially increasing data amount over time. The design relies on the use of CS-source coding instead of the supported speech coding standards in actual mobile communication systems. The proposed CS-source coding method (presented in Fig. 4.3) allows reducing the speech coding complexity by using simple

quantization and binary encoding, saving communication system resources, and encrypting communications without additional costs.

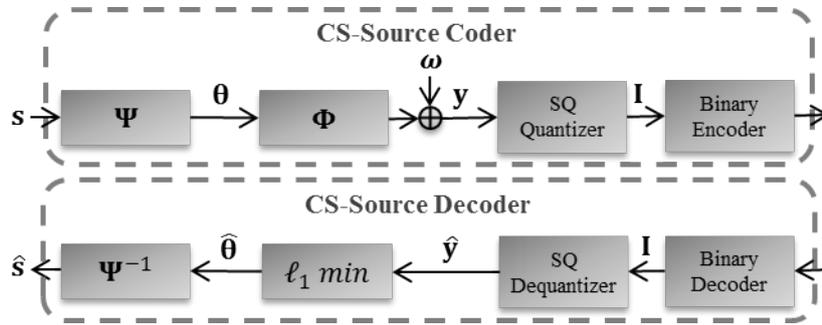


FIGURE 4.3: Proposed compressed sensing-based speech codec

In this work, we investigate the CS-source coding in real conditions by transmitting the compressed-sensed signals through a mobile communication system where multipath fading is always observed. Following, fading and channel modelling are clarified. Techniques chosen to compensate the channel and to ensure a robust communication are presented next.

4.3.2 Fading and Channel Modelling

The signals propagation over a mobile communication system is affected by three basic mechanisms: reflection, diffraction, and scattering. In fact, a signal may arrive at the receiver via multiple paths causing inter-symbol interference (ISI); this phenomenon is called multipath propagation. Another cause of ISI is the bandlimited character of many practical channels, for which the transmitted symbols tend to get "smeared" in time, and then interfered with. Mobile communication systems can also provide fluctuations in the amplitude and phase of the received signal, and the angle of arrival giving rise to multipath fading terminology (Sklar, 2001).

Multipath fading effects in mobile communication systems

Two types of fading effects characterize mobile communications: large-scale and small-scale fading (Sklar, 2001). Large-scale fading represents the average signal power attenuation or the path loss due to the motion over large areas. Small-scale fading refers to the dramatic changes in signal amplitude and phase that can be experienced as a result of small changes (as small as half wavelength) in the spatial positioning between a receiver and transmitter. In mobile systems, the channel is time varying due to the motion between the transmitter and the receiver, and this results in propagation paths changes. Two manifestations of small-scale fading are signal

time-dispersion (signal spreading) and time-variant nature of the channel (Sklar, 2001). For signal dispersion, the fading degradation can be frequency-selective or frequency-nonsselective (flat) fading. For the time-variant manifestation, the attenuation can be fast fading or slow fading. Let us define B_c as the minimum frequency distance where attenuations of the channel are uncorrelated, and T_c as the minimum temporal distance where the attenuations of channel are (nearly) constants. If the signal bandwidth is smaller than the coherence bandwidth B_c , then the fading is flat; otherwise, it is frequency-selective. If the signal duration is greater than the coherence time T_c , then fast fading occurs; otherwise, it is slow fading. The small scale fading (short-term) can be modelled as Rayleigh fading (Stüber, 2017).

Modelling the propagation channel

Mobile systems channels have multiple paths due to reflection, diffraction, and scattering caused by environment obstacles. Each path is characterized by an attenuation and a delay. The received signal $\mathbf{r}(n)$ is therefore a sum of translated versions, and it can be written as (Sklar, 2001):

$$\mathbf{r}(n) = \sum_k \mathbf{h}(n;k) \mathbf{a}(n-k) + \mathbf{w}(n) \quad (4.1)$$

where n denotes discrete time, $\mathbf{a}(n)$ is the transmitted symbol, $\mathbf{w}(n)$ is an additive white Gaussian noise (AWGN), and $\mathbf{h}(n;k)$ is the time varying channel impulse response. The impulse response of the channel $\mathbf{h}(n;k)$ is modelled by (Sklar, 2001):

$$\mathbf{h}(n;k) = |\mathbf{h}(n;k)| \exp(j\varphi(k)) \quad (4.2)$$

where $\varphi(k)$ represents a random variable phase uniformly distributed in $[-\pi, \pi]$, and the amplitude $|\mathbf{h}(n;k)|$ can follow several distributions such as Rayleigh or Rice models. In the case of non line of sight propagation, i.e., the received signal is the sum of only delayed signals due to multipath without having a direct path, then $|\mathbf{h}(n;k)|$ is a Rayleigh process. If the propagation is line of sight, i.e., there is a direct path providing a dominant component, then $|\mathbf{h}(n;k)|$ follows a Rice distribution (Stüber, 2017).

Rayleigh fading model is considered in this thesis as it is ideally suited to situations where there are large numbers of signal paths and reflections. It operates best under conditions when there is no dominant signal (direct line of sight signal), and in many instances cellular telephones being used in a dense urban environment fall into this category.

4.3.3 Multipath Fading Compensation

In a typical mobile radio propagation environment, the amplitude of the received signal is weakened or strengthened depending on the position of the terminal. Such communications suffer from ISI and signal attenuation due to the channel. Rayleigh channel is a statistical model of the fading channel recognized as the model of the worst case. It assumes that the power of a signal transmitted through such a communication channel varies randomly, or fades according to Rayleigh distribution.

The presence of ISI in the system makes the communication less reliable. It is an unwanted phenomenon as the interfering symbols have similar effect to noise, and this introduces errors in the decision device at the receiver output. Moreover, the attenuation, which is a reduction of signal strength during transmission, may significantly degrade the performance of a communication system. To mitigate these effects that are linked to the channel, various techniques are used. For instance, equalization, direct sequence spread spectrum, or OFDM can be applied to enhance the interfere-proof ability; and to cope with attenuation, different methods exist such as RAKE/Pre-RAKE receiver or exploiting diversity techniques. Brief definitions of these techniques are given below.

Anti-Multipath Systems

There are several signal processing solutions we can use as countermeasures to the ISI caused by delay spread. These techniques can be used at the transmitter or the receiver to make the signal less susceptible to delay spread. Equalization is a signal processing technique used at the receiver to alleviate the ISI problem. spread spectrum and multicarrier modulation fall in the category of transmitter signal processing techniques.

Equalization Equalization is the reversal of distortion incurred by a transmission channel. In other words, it is the process of reducing amplitude, frequency, and phase distortion due to ISI in a channel in order to improve transmission performance. Equalization is partitioned into two broad categories (Sklar, 2001). The first category is based on maximum-likelihood sequence estimation, and the second category uses filters to compensate the distorted pulses. The equalizers strengthen (boost) or weaken (cut) the energy of specific frequency bands. It must have an estimate of the channel impulse or frequency response¹. Since the wireless channel varies over time, the equalizer must learn the frequency response of the channel (training), and

¹As the channel estimation is out of our interest, in our thesis, the channel impulse response is considered known by the receiver

then update its estimate of the frequency response as the channel changes (tracking). This is often referred to as adaptive equalization. In general, the training is done by sending, over the channel, a fixed-length known bit sequence, which can be used by the equalizer at reception to estimate the channel frequency response.

Direct Sequence Spread Spectrum Spread spectrum techniques allow spreading the signal in the time or frequency domains. They reduce the risk of interference with other receivers while providing some privacy. One of the most used, direct sequence spread spectrum replaces “1” by a pseudo-random sequence and the “0” by its logical complement (chipping); thus, a redundancy is introduced for performing detection or even error correction at the end of transmission chain (Stüber, 2017).

OFDM Modulation is a multicarrier modulation technique consisting in transmitting digital data by modulating it on a large number of orthogonal carriers at the same time. The modulation and demodulation are implemented using circuits performing fast Fourier transforms. OFDM has been adopted in several wireless standards like 802.11a/g/n (WLAN), 802.16a (WiMAX), LTE, *etc.* (Stüber, 2017). It has been mainly designed to resist against the multipath fading and to ensure sufficiently high data rates. The advantages of OFDM are:

- Low ISI: Adding a cyclic prefix (CP) increases the robustness of the OFDM signal against multipath.
- Optimal spectral occupancy: Orthogonality between the n subcarriers allows overlapping their frequency bands, respectively, and thus optimizes the spectral efficiency of the modulated signal.
- Locally invariant Channel: The bandwidth of each subcarrier is small compared to the entire bandwidth of the OFDM signal. Consequently, it can be considered that the frequency response of the transmission channel is flat at each subcarrier. The frequency fading of the channel is, therefore, a “flat fading”.
- Simple frequency equalization: Complex equalization (inverse convolution) is reduced to a simple division in the frequency domain.

Hence, OFDM is used in this thesis for modulation.

Anti-Fading Systems

To counter the effects of multipath fading, path diversity can be exploited using RAKE/PreRAKE receiver. In addition, different other diversity schemes can be used.

RAKE/PreRAKE Receiver A RAKE receiver is a radio receiver that overcomes the fading due to multipath phenomena. It consists of several sub-receptors called “fingers”. Each finger is a correlator associated with a multipath propagation component. It independently decodes the multipath components. The contribution of each sub-receptor is then recombined to coherently sum all these multiple paths. This technique improves the SNR of a transmission in a dense multipath propagation channel (such as Rayleigh channel). However, the RAKE receiver requires some additional signal processing for setting the weighting factors and the combination functions. It causes more complexity and higher energy consumption in the receiver. In mobile communications, it is desirable to reduce power consumption, size, and cost of a mobile terminal in the downlink. PreRAKE was proposed to transfer the receiver signal processing load to the transmitter (Stüber, 2017).

Diversity Technique Even if diversity is a form of redundancy, it is considered as a useful solution to the problem of multipath fading in wireless communications. Its principle is to provide for the receiver several versions of the same signal on independent channels. These versions are unlikely to fade simultaneously. There are mainly three types of diversity which are (Stüber, 2017):

1. frequency diversity;
2. time diversity (signal repetition); and
3. space diversity.

Frequency diversity uses multiple carriers separated by Δf greater than the coherence bandwidth B_c of the channel. The system can then provide independent versions of fading. This can be achieved, e.g., by frequency hopping, or spreading matrices. The disadvantage of this method is its low spectral efficiency.

By transmitting the signal in different time slots, the system can achieve temporal diversity. The interval between successive time slots must be greater than or equal to the coherence time T_c ; otherwise, the system performance degrades. This can be achieved by coding and interleaving. Its disadvantage is reduced system capacity.

To provide independent fading, one can exploit space diversity using multiple transmit and/or receive antennas, and choosing the spacing between the antennas ($> \lambda/2$) (Stüber, 2017).

Depending on the antenna combination in the transmitter or the receiver, one can class space diversity as:

1. receive diversity, which uses a single transmitting antenna and multiple receiving antennas (SIMO: Single Input Multiple Output);
2. transmit diversity, which consists of using multiple transmitting antennas and a single receiving antenna (MISO: Multiple Input Single Output); or
3. mixed diversity in transmission and reception, which combines the use of both multiple transmitting and multiple receiving antennas (MIMO: Multiple Input Multiple Output).

In this work, space diversity, using multiple receiving antenna, is used at the receiver (SIMO).

The main objective of SIMO systems is to reduce the amplitude fluctuations due to fading. It uses signals received by two or more uncorrelated antennas. If an antenna suffers a fading, it may not reach the other antennas, and at least, one good signal can be received.

The various received versions are combined at the reception so that the fading is reduced. Three processing techniques can be used to combine diversity (Stüber, 2017): Selection combining (SC), equal gain combining (EGC), and maximum ratio combining (MRC). In SC, the idea is to select the antenna with the best signal (usually the signal strength is taken as a measure of signal quality). EGC seeks to improve results by co-phasing signals (by applying the same weighting factor), and adding them. Maximum ratio combining is the most optimal method. It differs from EGC by using different weighting factors dependent on the SNRs. For the proposed communication system design the three combiners are, firstly, assessed to choose the best for our system.

In addition to equalization, OFDM modulation, and diversity, channel coding is used for error detection and correction as described in the next subsection.

4.3.4 Channel Coding

Channel coding is used, in communication systems (Fig. 4.1), to minimize error rate. The principle is to add additional digits to the transmitted information. The added digits do not carry new information, but they allow the receiver detecting and correcting errors. Error correcting codes are divided into two main classes: block codes and convolutional codes. In convolutional coding, the bits are processed serially and not in blocks. The redundant bits are generated by the use of modulo-2 convolutions in the encoder; hence the name. A convolutional encoder with k inputs and n outputs is of rate k/n . The convolutional code is specified by the parameters

(n,k,L) , where L is the constraint length of the code. Convolutional codes are applied in almost all communication systems after the apparition of Viterbi's algorithm for decoding (Ruyet and Pischella, 2015). Therefore, convolutional coding and Viterbi decoding are used in this work for error detection/correction. More details about channel coding can be found in (Stüber, 2017; Ruyet and Pischella, 2015).

4.4 Improving the System for Lower Bitrates

The proposed communication system integrate a CS-based speech codec that exploits scalar quantization (SQ) for CS measurements quantization. Even if this codec is very effective in terms of quality, intelligibility, and execution time, as discussed in the previous chapter, it can not achieve very low bitrates. To improve the proposed communication system for lower bitrates, we propose the use of CS speech codec (see Fig. 4.4) that uses split-multistage vector quantization (S-MSVQ) instead of SQ.

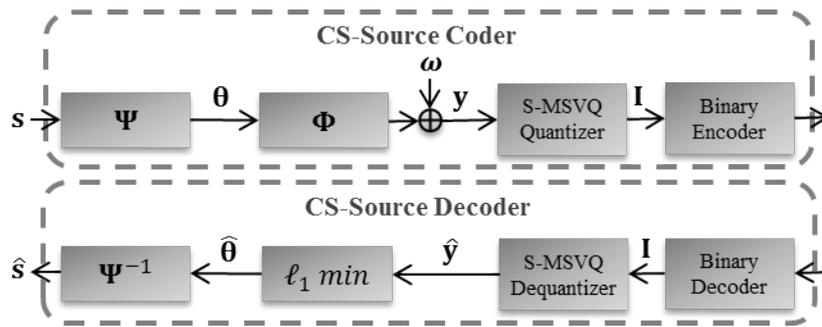


FIGURE 4.4: Proposed compressed sensing-based speech codec

The evaluation of the performance of the proposed communication system and its improved version is performed and presented in the following section.

4.5 Experiments

The proposed end-to-end transmission scheme (Fig. 4.5) increases the transmission rate by using CS-source coding for speech compression in the context of mobile communication systems. After applying CS, the obtained signal, called observation vector, is quantized and binary encoded. In order to minimize error rate, channel encoding is performed using convolutional code. The obtained data are then modulated using OFDM and transmitted over a Rayleigh multipath channel. The signal is received at multiple antennas (receive diversity). OFDM demodulation, diversity combining, and equalization are, then, applied. In this work, the three combining techniques (SC, EGC, and MRC) are tested and compared to select the best one

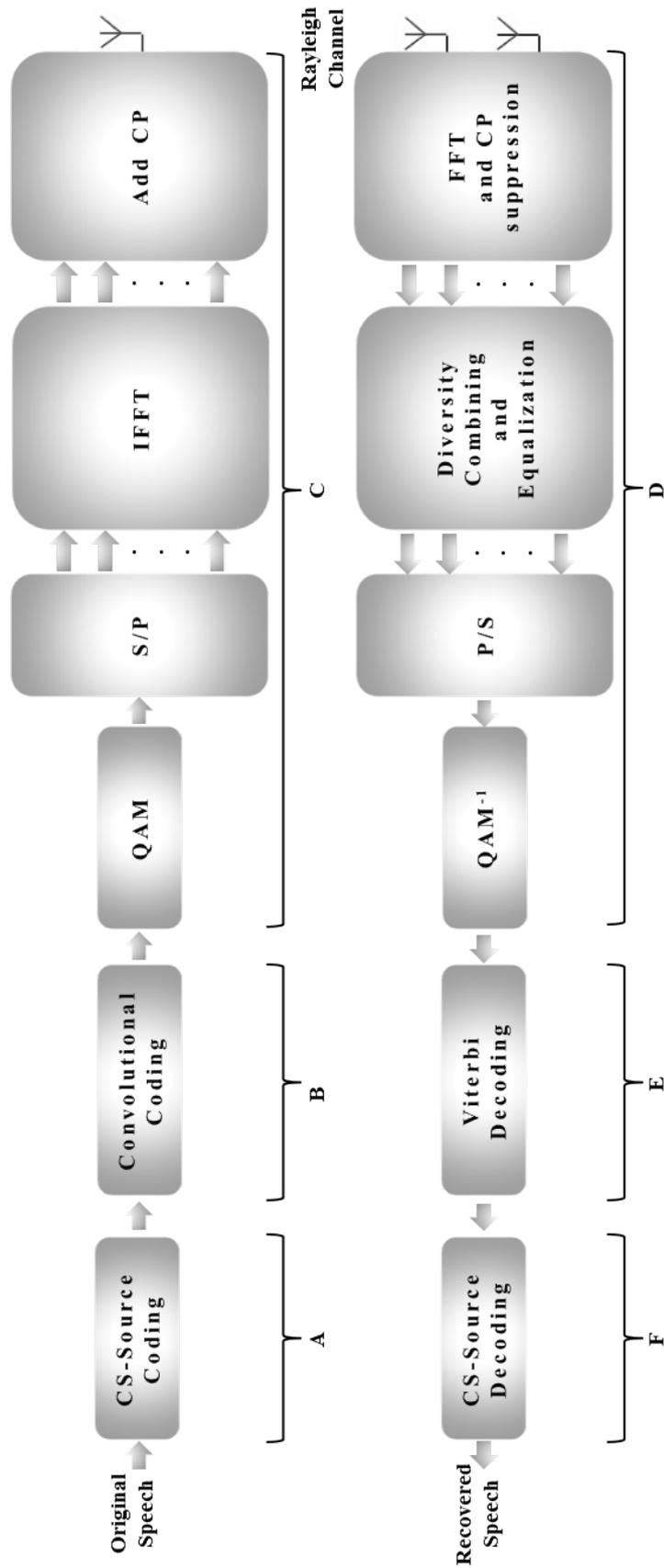


FIGURE 4.5: Mobile communication system using compressed sensing.

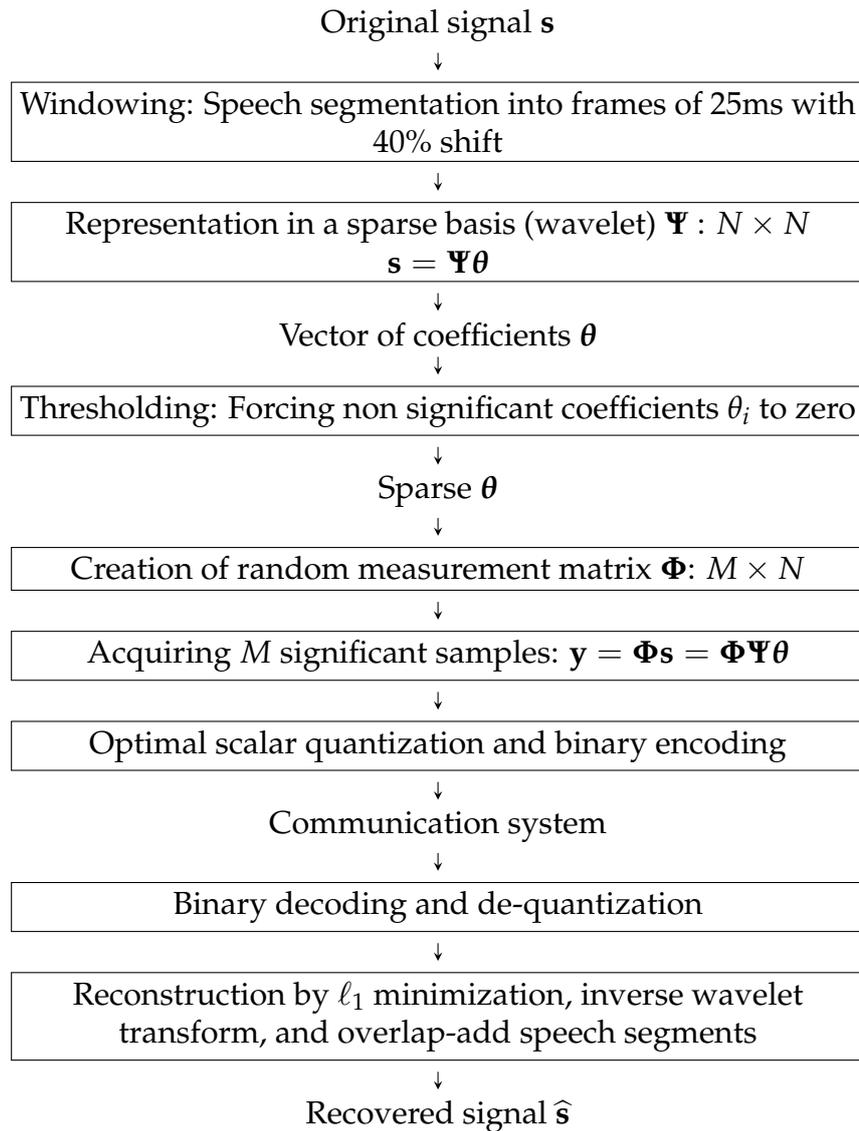


FIGURE 4.6: Flowchart of compressed sensing-source coding.

regarding the bit error rate (BER) as will be seen. After that channel decoding using Viterbi algorithm is performed. To recover the speech signal, CS-source decoding is performed by: binary decoding, dequantization, and CS reconstruction. Fig. 4.5 shows the steps of the simulation.

In the proposed system, the speech signal is, directly, source-coded using a method based on CS (block A). Fig. 4.6 illustrates the CS-source coding implementation steps when SQ is considered.

The speech signal is, firstly, segmented. The sparse representation is obtained by projecting each speech segment in the wavelet basis (Ouahabi, 2012), and then, performing the thresholding (Algorithm 4). As the energy is concentrated in the

coefficients of higher amplitude, eliminating the lower coefficients will improve the sparsity without involving a significant signal loss. This can be done by applying hard thresholding using the universal threshold γ defined as (Donoho and Johnstone, 1994; Donoho, 1995):

$$\gamma = \sigma \sqrt{2 \log_e N} \quad (4.3)$$

where N is the length of the vector of coefficients, and σ^2 is the noise variance. σ is estimated by $\hat{\sigma} = m/0.6745$, where m is the median of wavelet coefficients at level one. After zeroing the wavelet coefficients (i.e., hard thresholding) greater than the threshold γ , an $M \times N$ Gaussian random measurement matrix is created. This matrix fulfills the RIP and incoherence conditions. M random values are acquired by multiplying the sparse vector by this matrix. The result (observation vector) is a compressed and secured signal, due to the randomness of the measurement matrix that encrypt the signal. This encryption is another advantage of using a random measurement matrix. The observation vector is then quantized using optimal SQ (using Lloyd optimization)², and then binary encoded.

Algorithm 4 Performing compressed sensing on speech

Input: \mathbf{s} the input speech signal and M the number of measurements

Output: The observation vector \mathbf{y}

```

1: function COMPRESSED SENSING( $\mathbf{s}$ ,  $M$ )
2:   Sparse representation
3:    $\boldsymbol{\theta} \leftarrow \boldsymbol{\Psi} \mathbf{s}$  ▷  $\boldsymbol{\Psi}$ : "db1" wavelet basis
4:   Compute threshold  $\gamma$ 
5:    $\sigma \leftarrow \text{median}(\boldsymbol{\theta}) / 0.6745$ 
6:    $N \leftarrow \text{length}(\boldsymbol{\theta})$ 
7:    $\gamma \leftarrow \sigma \sqrt{2 \log_e N}$ 
8:   Perform thresholding
9:   for  $i \leftarrow 1$  to  $N$  do
10:     if  $|\theta(i)| < \gamma$  then
11:        $\theta(i) \leftarrow 0$  ▷ Hard thresholding
12:     end if
13:   end for
14:   Create random measurement matrix  $\boldsymbol{\Phi}$  ▷ of length  $M \times N$ 
15:   Perform acquisition
16:    $\mathbf{y} \leftarrow \boldsymbol{\Phi} \boldsymbol{\theta}$  ▷  $\mathbf{y}$  has length  $M$ 
17:   return  $\mathbf{y}$ 
18: end function

```

The compressed sensing-source coded signal is channel encoded (block B) using 1/2 rate convolutional code (the constraint length is equal to 7 and the code generator

²For the improved version of the proposed communication system, split-multistage vector quantization is used in the simulation.

is [171 133]). OFDM modulation is then performed, as seen in Fig. 4.5 (block C). In OFDM (Algorithm 5), the signal bandwidth B_s is divided into L sub-bands of Δf bandwidth, with parameters chosen such as

$$\Delta f \ll B_c \quad (4.4)$$

and

$$T_s \ll T_c \quad (4.5)$$

where T_s is the symbol period, B_c is the coherence bandwidth of the channel, and T_c is the coherence time of the channel. This results in a frequency non-selective (in the band Δf of subcarrier) and time-invariant (during the period T_s of a symbol) channel. The insertion of the cyclic prefix (CP) allows spreading the duration of an OFDM symbol to prevent interference between OFDM symbols (ISI).

Algorithm 5 Performing OFDM

Input: \mathbf{y} the signal to transmit, Nb the OFDM symbol length, and CP the cyclic prefix duration

Output: The modulated symbols \mathbf{Y}_{ifftCP}

```

1: function OFDM( $\mathbf{y}$ ,  $Nb$ ,  $CP$ )
2:   Perform QAM-16 modulation
3:    $\mathbf{y}_{qam} \leftarrow \text{QAM16}(\mathbf{y})$ 
4:   Form  $NbSymb$  symbols Symb of length  $Nb$ 
5:    $NbSymb \leftarrow \text{length}(\mathbf{y}_{qam}) / Nb$ 
6:   Perform IFFT for each symbol
7:   for  $i \leftarrow 1$  to  $NbSymb$  do
8:      $\mathbf{Y}_{ifft} \leftarrow \text{ifft}(\mathbf{Symb}(i))$ 
9:     Add cyclic prefix to get  $\mathbf{Y}_{ifftCP}$  ▷ Length  $\mathbf{Y}_{ifftCP}$  is  $Nb + CP$ 
10:  end for
11:  return  $NbSymb$  OFDM symbols ( $\mathbf{Y}_{ifftCP}$ )
12: end function
```

For OFDM, the quadrature amplitude modulation (QAM) is applied to the observation vector. Then, the serial to parallel conversion (S/P) and inverse Fourier transform are performed, and the CP is added. The principle of QAM is to form p -bit symbols (p integer), and then represent them in a constellation. A constellation may be of different sizes depending on the desired number of bits per symbol. Indeed, a constellation represents all the possible states of these symbols. Thus, if we want p -bit symbols, we will have 2^p possible symbols corresponding to 2^p points in the constellation, arranged in a square grid with equal vertical and horizontal spacing. By using more points on the constellation, it is possible to transmit more bits per symbol. However the points are closer together and they are therefore more susceptible to noise and data errors. In this simulation we used 16-QAM, i.e., a

constellation of 16 symbols, as a compromise between transmission rate and error. Let $D_p = A_p \exp(j\vartheta_p)$ be the symbol associated to the 16-QAM constellation point, where A_p is its amplitude, and ϑ_p is its phase. Since all sub-carriers are transmitted in parallel, the OFDM symbol $\mathbf{a}(n)$ can be written as:

$$\mathbf{a}(n) = \sum_{p=1}^N D_p \exp(j2\pi pn/N) \quad (4.6)$$

where D_p is the constellation point, n is a discrete time variable. Equation (4.6) can be thought of as an IFFT process, where N is the IFFT size. To add a CP, one can simply copy the end of the OFDM symbol to which it is added.

The OFDM symbols are transmitted through the Rayleigh multipath channel. This simulation assumes that the number of taps in the channel is lower than the CP duration (to reduce ISI). Since the defined CP is of eight samples, the Rayleigh channel is chosen to be of seven taps. This means that the received signal is the sum of seven delayed versions of the transmitted signal. By choosing a relevant OFDM symbol duration ($T_s \ll T_c$), one can obtain a time invariant channel during the symbol period; hence the received symbol is expressed as:

$$\mathbf{r}(n) = \mathbf{h}(n) * \mathbf{a}(n) + \mathbf{w}(n) \quad (4.7)$$

where n denotes a discrete time, $\mathbf{r}(n)$ is the received symbol, $\mathbf{h}(n)$ is the channel impulse response statistically described by a Rayleigh distribution, $\mathbf{a}(n)$ is the transmitted symbol, $\mathbf{w}(n)$ is an AWGN, and $*$ denotes the convolution operator.

The propagation in such a channel causes ISI and attenuation. OFDM overtakes the ISI between symbols thanks to CP insertion. To compensate signal attenuation, we used the technique of spatial receive diversity by utilizing multiple antennas in reception.

After the CP extraction, the received symbols (by the various antennas) are transformed to the frequency domain. They are, next, combined using SC (Algorithm 6), EGC (Algorithm 7), or MRC (Algorithm 8). In the next step, the equalization, parallel to serial conversion (P/S), and the QAM demodulation are performed (block D of Fig. 4.5).

Channel decoding is performed using Viterbi algorithm (block E). In CS-source decoding (block F), binary decoding is, firstly, performed. The result is, secondly, de-quantized to restore the observation vector. This vector passes finally to the reconstruction module which recovers the original signal by performing ℓ_1 minimization,

Algorithm 6 Selection combining algorithm

Input: **SymbF** a matrix of length $nRx \times Nb$ that contains the frequency domain received OFDM symbols in lines, and **h** a matrix of length $nRx \times nTap$ where each line is the channel response of a path

Output: The combined OFDM symbol **SymbF_{hat}**

```

1: function SELECTION COMBINING(SymbF, h)
2:   Compute the frequency response of the channels H
3:   H  $\leftarrow$  fft(h)
4:   Compute the channels power Hpower
5:   Hpower  $\leftarrow$  H * conj(H)
6:   Find the maximum power Hpowermax
7:   Hpowermax  $\leftarrow$  max(Hpower)
8:   Select the chain with the maximum power
9:   Hsel  $\leftarrow$  H(Hpower == Hpowermax)
10:  SymbFsel  $\leftarrow$  SymbF(Hpower == Hpowermax)
11:  Equalization with the selected chain
12:  SymbFhat  $\leftarrow$  SymbFsel ./ Hsel
13:  return SymbFhat
14: end function

```

Algorithm 7 Equal gain combining algorithm

Input: **SymbF** a matrix of length $nRx \times Nb$ that contains the frequency domain received OFDM symbols in lines, and **h** a matrix of length $nRx \times nTap$ where each line is the channel response of a path

Output: The combined OFDM symbol **SymbF_{hat}**

```

1: function EQUAL GAIN COMBINING(SymbF, h)
2:   Compute the frequency response of the channels H
3:   H  $\leftarrow$  fft(h)
4:   Divide by the phase of H
5:   SymbF  $\leftarrow$  SymbF * exp(-j * angle(H))
6:   Compute the sum
7:   SymbFhat  $\leftarrow$  sum(SymbF, 1) ./ sum(abs(H), 1)
8:   return SymbFhat
9: end function

```

Algorithm 8 Maximal ratio combining algorithm

Input: \mathbf{SymbF} a matrix of length $nRx \times Nb$ that contains the frequency domain received OFDM symbols in lines, and \mathbf{h} a matrix of length $nRx \times nTap$ where each line is the channel response of a path

Output: The combined OFDM symbol \mathbf{SymbF}_{hat}

```

1: function MAXIMAL_RATIO_COMBINING( $\mathbf{SymbF}$ ,  $\mathbf{h}$ )
2:   Compute the frequency response of the channels  $\mathbf{H}$ 
3:    $\mathbf{H} \leftarrow fft(\mathbf{h})$ 
4:   Compute the combined signal
5:    $\mathbf{SymbF}_{hat} \leftarrow \frac{sum(conj(\mathbf{H}) .* \mathbf{SymbF}, 1).}{sum(\mathbf{H} .* conj(\mathbf{H}), 1)}$ 
6:   return  $\mathbf{SymbF}_{hat}$ 
7: end function

```

inverse wavelet transform, and overlap-add speech segments. We consider basis pursuit because:

1. minimizing the l_1 norm of a solution promotes sparsity,
2. the sparse approximation problem can be replaced by a convex problem, and there are many computationally efficient algorithms that can find the solutions, and
3. l_1 minimization is known by its robustness.

The parameters of the simulation and performance measures are presented next.

4.5.1 Simulation Setup and Performance Measures

A new mobile communication system that integrates CS-speech codec instead of the current speech codecs (e.g., AMR-WB) is designed. Simulations were carried out to evaluate the performance of

- (i) the proposed communication system when the integrated CS-speech codec is based on SQ of CS measurements; and
- (ii) its improved version that considers a CS-speech codec based on S-MSVQ of CS measurements.

The experiments used clean speech signals extracted from TIMIT database (Garofolo et al., 1993). The sampling frequency is $f_s = 16$ kHz. A Hamming window was used for speech segmentation. The frame size was chosen to be 25 ms with a 40% shift. This test was conducted using MATLAB. Different compression ratios (C) are considered when performing CS.

For SQ, the number of bits per sample ($nbps$) was chosen to be 4 based on the results of the previous chapter. For S-MSVQ, the codebooks were created using a training database composed of CS measurements (observation vectors) obtained by performing CS on the training speech signals at different values of C . Thus, we have a different training database for each C . Note that 50 utterances were considered for training, and 20 utterances, different from the latter, were used for test (of both genders).

The parameters of the simulation are summarized in Table 4.1.

TABLE 4.1: Simulation parameters

Parameter	Value
Sparse basis	Wavelet ($db1$ was used)
Measurement matrix	Gaussian random
Channel coding	1/2 convolutional code
Channel decoding	Viterbi algorithm
OFDM	QAM16, $nFFT = 64$, $CP = 1/8$
Channel model	Rayleigh
Number of paths	$nTaps = 7$
Diversity technique	Receive-antenna diversity, $nRx = 1$ to 3
Combining techniques	SC, EGC and MRC
Recovery method	ℓ_1 minimization

The performances of the proposed communication systems are evaluated in terms of wideband extension of the perceptual evaluation of speech quality (WB-PESQ) and speech intelligibility index (SII). Refer to Section 3.7.2 for more details about these measures.

4.5.2 Results

Thanks to the sparsity of speech in the wavelet basis, CS can be used for simultaneously acquiring and compressing signals. Exploiting this paradigm, besides other signal processing techniques, for speech communication has resulted in a new design for mobile systems.

Evaluation of the Proposed Communication System

Fig. 4.7 shows a zoom of 1000 samples of an example of simulation as well as its wavelet coefficients after thresholding.

Figs. 4.8, 4.9, and 4.10 present the variation of BER *vs.* E_b/N_0 when performing the three diversity combining techniques (SC, EGC, and MRC respectively) with

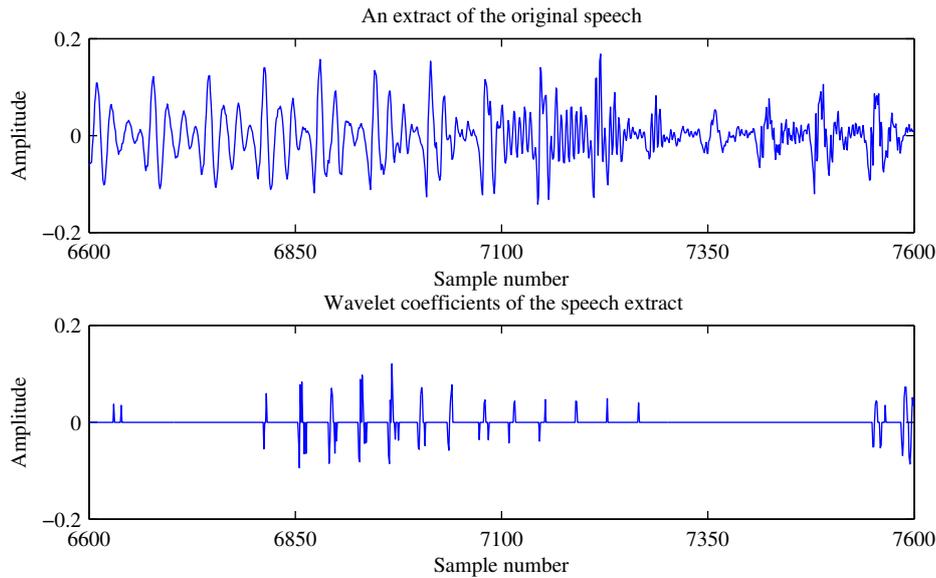


FIGURE 4.7: An extract of 1000 samples of original speech signal (top), and its sparse representation (bottom). (Observe that the wavelet representation is sparser than the temporal one).

different numbers of antennas in reception. As expected, they show that the BER

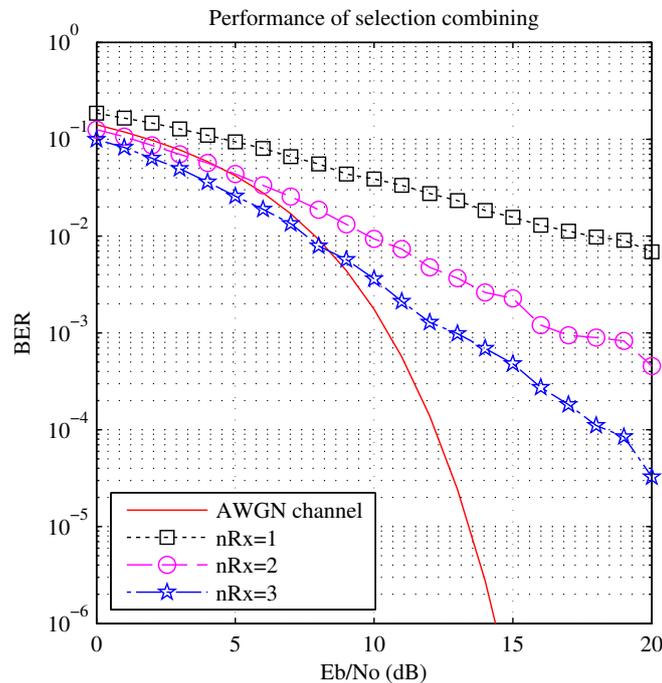


FIGURE 4.8: Performance of selection combining diversity with OFDM modulation in Rayleigh channel. (Observe that with SC the BERs are reduced when the number of receiving antennas increases: Around 8 dB improvement at 10^{-2} BER point with $nRx = 2$ and 10 dB with $nRx = 3$).

decreases with the rising of the number of receiving antennas. The more the number of receiving antennas increases, the more the performances approach those of an

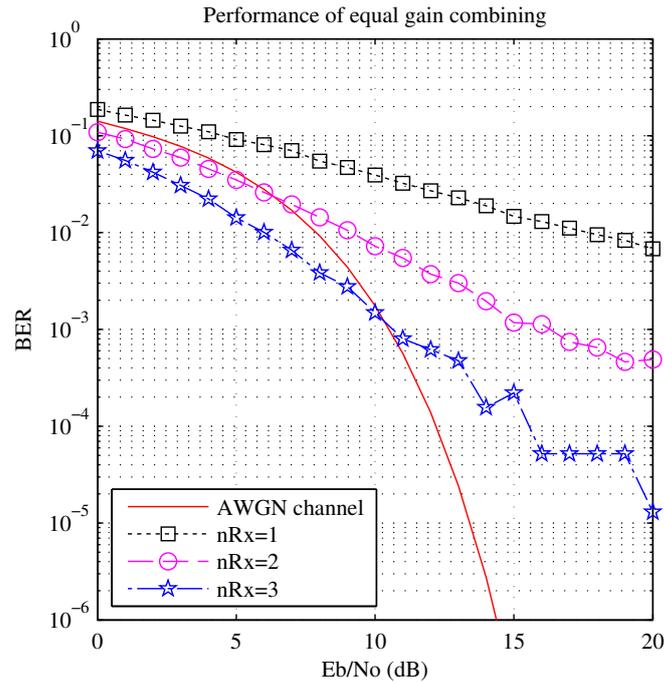


FIGURE 4.9: Performance of equal gain diversity combining with OFDM modulation in Rayleigh channel. (Observe that with EGC the BERs are reduced when the number of receiving antennas increases: Around 9 dB improvement at 10^{-2} BER point with $n_{Rx} = 2$ and 12 dB with $n_{Rx} = 3$).

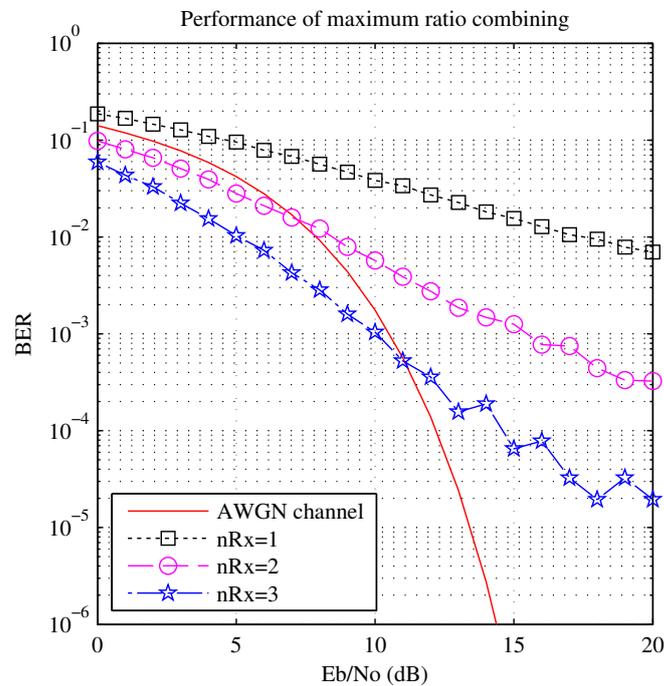


FIGURE 4.10: Performance of maximum ratio diversity combining with OFDM modulation in Rayleigh channel. (Observe that with MRC the BERs are reduced when the number of receiving antennas increases: Around 10 dB improvement at 10^{-2} BER point with $n_{Rx} = 2$ and 13 dB with $n_{Rx} = 3$).

AWGN channel in which the fading effects (in Eq. (4.7)) are not considered, i.e., $\mathbf{r}(n) = \mathbf{a}(n) + \mathbf{w}(n)$. In real systems, the number of receiving antennas is limited to the constraint of mobile phone size (e.g., two antennas are used in the current mobile devices: The main antenna and the diversity antenna); therefore, we have limited our study to two receiving antennas while three receiving antennas are used to show the effect of increasing the number of antennas on the BER. However, this study can be extended to more antennas to accommodate future 5G systems. Fig. 4.11 compares the BERs of the three techniques. The BERs of MRC combiner are smaller than those of EGC and SC.

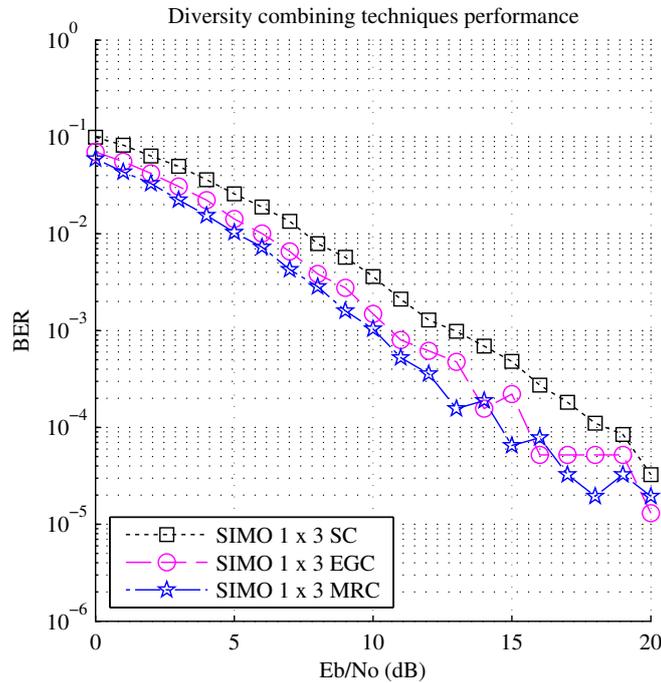


FIGURE 4.11: Performance comparison of diversity combining techniques with OFDM modulation in Rayleigh channel ($nR_x=3$). (Observe that the BERs of MRC are the smallest).

Fig. 4.12 shows the transmitted constellation and the received constellation after using two receiving antennas and MRC combiner with a compression ratio $C = 20\%$ and $nbps = 4$ ($BR = 12.8$ kbit/s). One can observe that the constellation is recovered but with few grains.

Fig. 4.13 presents the original and the recovered speech signals after performing channel decoding and CS-source decoding. To quantify the difference between the two signals, we computed the MSE. For this example, an $MSE = 1.2 \times 10^{-3}$ and a $PESQ = 3.43$ are obtained.

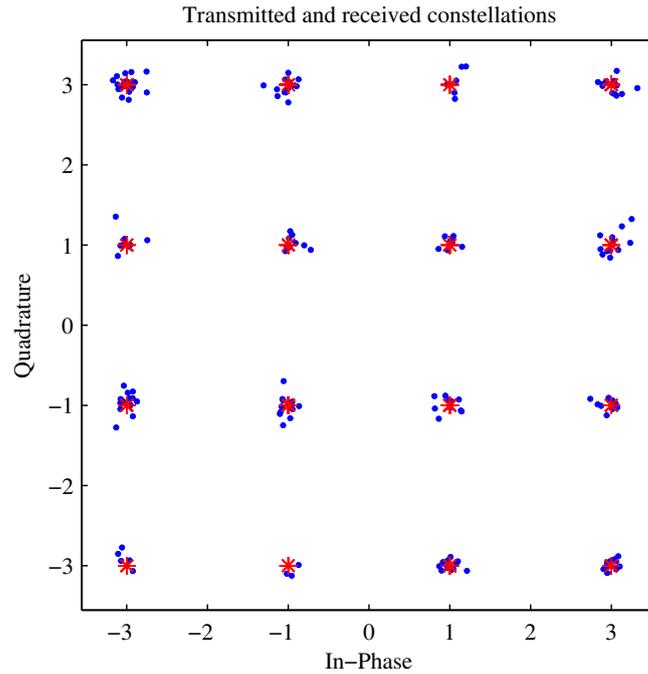


FIGURE 4.12: Constellation diagrams of the transmitted signal (in red asterisks) and the received signal (in blue dots) after equalization with receive diversity ($BR = 12.8$ kbit/s; $nR_x=2$; MRC; $E_b/N_0= 20$ dB). (Observe that the QAM-16 constellation is recovered with few distortion due to the white Gaussian noise).

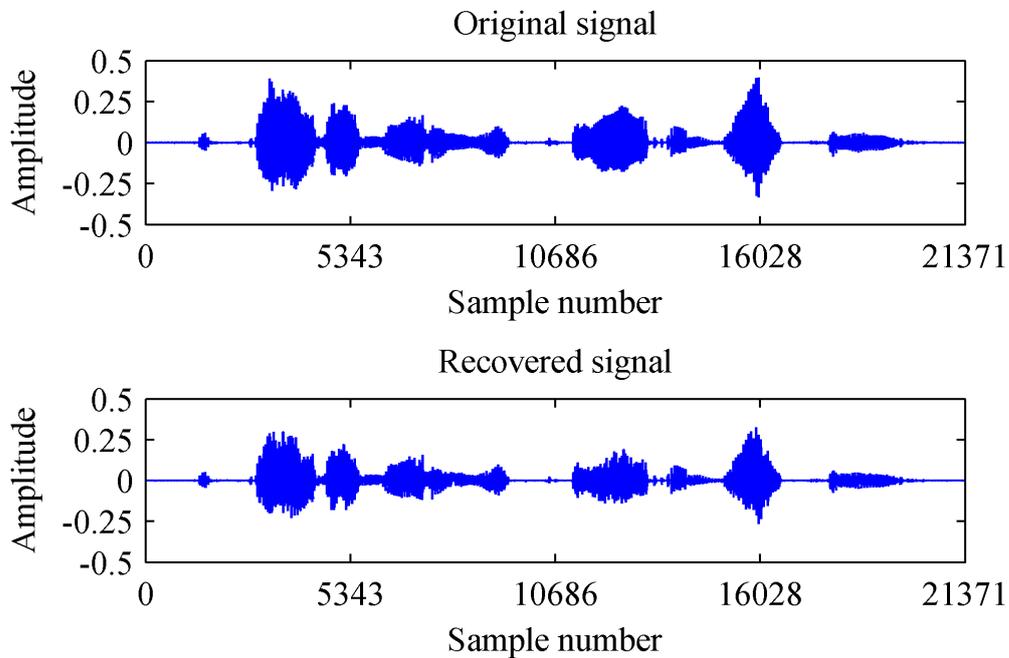


FIGURE 4.13: The original (top) and the recovered (bottom) speech signals. (Note that the mean-squared error of this example is 1.2×10^{-3} , and the PESQ is 3.43).

Table 4.2 lists the average PESQ scores of the speech issued from the proposed communication system (that integrates CS-source coding). It also clarifies the effects of the channel ($nRx = 1$), the diversity ($nRx = 2$), and the compression ratio ($C = M/N$) on the quality of output speech. Here, the number of antennas is limited to 2 (as in actual mobile phones) and $nbps = 4$ in 10 dB environment. The average SII values of the proposed chain are presented in Table 4.3.

TABLE 4.2: The average PESQ scores of the recovered speech signals ($nbps = 4$, $E_b/N_0 = 10$ dB).

C (%)	BR (kbit/s)	PESQ	
		$nRx = 1$	$nRx = 2$
10	6.4	2.44	2.39
20	12.8	2.86	3.33
30	19.2	2.77	2.81
40	25.6	2.88	2.99
50	32	3.41	3.80
60	38.4	3.12	3.82
70	44.8	3.00	3.70
80	51.2	3.35	3.73
90	57.6	3.63	3.74

TABLE 4.3: The average SII values of the recovered speech signals ($nbps = 4$, $E_b/N_0 = 10$ dB).

C (%)	BR (kbit/s)	SII	
		$nRx = 1$	$nRx = 2$
10	6.4	0.42	0.41
20	12.8	0.45	0.50
30	19.2	0.50	0.53
40	25.6	0.57	0.61
50	32	0.61	0.72
60	38.4	0.69	0.83
70	44.8	0.70	0.85
80	51.2	0.73	0.86
90	57.6	0.76	0.85

Evaluation of the improved version of the Proposed Communication System

An improved version of the proposed communication system is designed for low bitrates. It was evaluated for $BR = 8.85$ kbit/s using different number of stages/splits and various values of C . We chose this bitrate because it is one of the nine BR s (modes) of AMR-WB which is used in GSM and UMTS networks. We did not

considered higher BR s in the evaluation of the improved version because of the memory constraints.

The average quality and intelligibility scores are presented in Figs. 4.14 and 4.15, respectively. One can observe the increase of SII and PESQ scores with the rising of C .

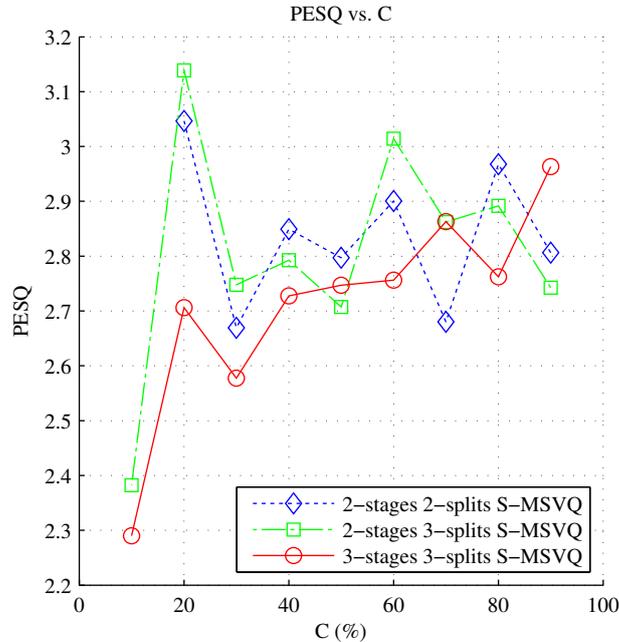


FIGURE 4.14: PESQ performances of the reconstructed signal for the proposed communication system, $BR= 8.85$ kbit/s. (Observe that the best quality is obtained when using 2-stages 3-splits S-MSVQ for a $CR=20\%$)

Table 4.4 lists the average PESQ and SII scores of the recovered speech from the proposed mobile communication system and its improved version. It also highlights a comparison with the PESQ results of some other CS-based speech compression methods.

4.5.3 Discussion

Performing both acquisition and compression simultaneously, by CS, reduces the processing burden at the transmitter. Hence, less pre-processing time and communication resources are needed, e.g., less battery consumption is required compared to current speech coding standards, such as AMR-WB that allows good quality at the cost of high battery consumption. By decreasing the amount of data, the time required for transmission across the network is reduced. Consequently, the transmission speed is increased. Because the condensed data take up less bandwidth, one can transmit greater volumes at a time, and this increases the transmission rate. In

TABLE 4.4: Performance of the proposed communication system and its improved version for $BR = 8.85$ kbit/s

Method	Compression Ratio	Channel	PESQ	SII
Proposed	13.83% ($BR = 8.85$ kbit/s)	10 dB Rayleigh	2.41	0.45
Improved	20% ($BR = 8.85$ kbit/s)	10 dB Rayleigh	3.14	0.47
(Giacobello et al., 2012)	50% ($BR = 10.1$ kbit/s)	Noise-free 10 dB AWGN	4.13 ± 0.13 3.21 ± 0.19	– –
(Al-Azawi and Gaze, 2018)	40% 60%	Noise-free	2.73 4.43	– –
(Ji, Zhu, and Champagne, 2019)	30% 80%	AWGN ($\mu = 0, \sigma = 0.002$)	2.40 (male) / 2.25 (female) 2.71 (male) / 2.80 (female)	– –

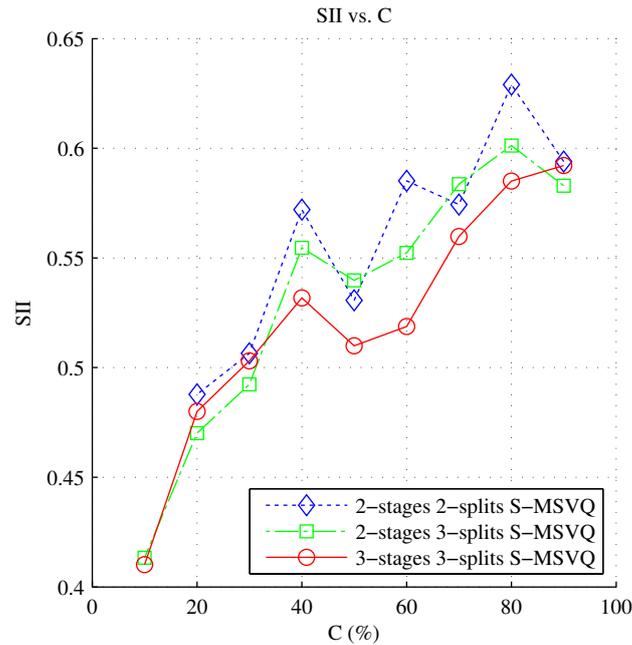


FIGURE 4.15: Intelligibility performances of reconstructed signal for the proposed communication system, BR= 8.85 kbit/s. (Observe that the intelligibility score increases with the rising of CR)

addition, the randomness of the measurement matrix may enhance the security of the communication because it can be considered as a form of encryption (Cambareri et al., 2015). The results presented here show that the proposed communication system design allows promising performances considering real transmission conditions.

Quality and Intelligibility Assessment

The use of CS for speech compression is possible thanks to the sparseness of speech in the wavelet domain as evidenced in Fig. 4.7. The latter shows that the wavelet coefficients present more null values; hence the sparsity.

The compressed signal is convolutionally encoded to increase the robustness against error, and modulated with OFDM to enhance the spectral efficiency and also the immunity against ISI caused by Rayleigh multipath fading channel.

The transmitted symbols, presented by the red asterisks in the constellation diagram in Fig. 4.12, are distorted after passing through this channel. To reduce this distortion, receive antenna diversity is used.

With the tested combiners: SC, EGC, and MRC (Figs. 4.8, 4.9, and 4.10), one notices that when the number of antennas increases, the BER *vs.* E_b/N_0 decreases faster. For instance, with MRC and for a BER of 10^{-2} , E_b/N_0 gains of 10 dB and 13 dB are obtained for two and three receiving antennas respectively. So, using more

receiving antennas leads to higher quality. This means that the diversity (within the proposed communication scheme) keeps its performance, and therefore it can be used in systems based on CS to mitigate the attenuation. Results in Table 4.2 support this deduction. Improved PESQ scores are obtained using two antennas with MRC combiner, e.g., for $C = 20\%$, the average PESQ score is 2.86. It increases to 3.33 when using two receiving antennas; hence the channel compensation.

Analysing Fig. 4.11, one notices that MRC gives the smallest BER comparing to EGC and SC; for this reason, MRC is adopted in the proposed system.

Fig. 4.12 shows that with two receiving antennas and MRC combiner, one gets the transmitted constellation but with some grains (the blue dots in the diagram) due to the residual AWGN (see Eq. (4.7)).

Fig. 4.13 indicates that the two signals are almost similar. An $MSE = 1.2 \times 10^{-3}$ and a $PESQ = 3.43$ are reported for this example, demonstrating a successful recovery. In fact, our communication system provides a good quality of recovered speech justified by a PESQ score approaching 3.5 for $BR = 12.8$ kbit/s. One can justify this by the robustness of ℓ_1 minimization, and the use of wavelet thresholding (performed to enhance the sparsity) which is considered as a denoising method in many applications.

Let us remind that a $PESQ \geq 2.5$ is interpreted as a *good* quality score, and a $SII \geq 0.75$ is a *good* intelligibility score. Note that this communication system was investigated for real communication conditions (10 dB Rayleigh channel). The promising PESQ values presented in Table 4.2 prove that the techniques used for mitigating channel effects allow channel compensation. One also notices from Table 4.2 (respectively Table 4.3) the improvement of PESQ scores (respectively SII values) of the output speech of the proposed system with the raising of compression ratio ($C = M/N$). With the same BR (12.8 kbit/s corresponding to $C = 20\%$), our CS-based communication system achieves *fair* intelligibility score of $SII = 0.50$ (Table 4.3). *Good* SII values are observed for high compression ratios.

Our CS-source coding scheme was compared with AMR-WB speech encoder, which is used in current mobile communication systems. At a $BR = 12.65$ kbit/s and for a channel error probability $p_e = 0.01$, the average PESQ score obtained using the original AMR-WB with 46 bits/frame split-multistage vector quantization is 2.86, and it is 1.35 for a $p_e = 0.1$ (Cheraitia and Bouzid, 2014, Table 8). With a $BR = 12.64$ kbit/s (corresponding to $C = 19.75\%$), our CS-source coding scheme achieves an average PESQ score of 2.40 for a $p_e = 0.01$, and 2.33 for a $p_e = 0.1$. One can notice that for more dramatic communications conditions (higher channel error

probability, i.e., $p_e = 0.1$), the PESQ results reveal the advantages of the proposed design. Furthermore, a gain in processing time performance is achieved. The latter is justified by the reduced complexity of source-coding and the amount of data, that decreases the processing burden in all the steps of the communication.

The improved version of the proposed communication systems also shows promising results. Considering a $BR = 8.85$ kbit/s, *fair* intelligibility scores are achieved for all the studied values of C . It is also noticed that the results obtained using two-stages S-MSVQ quantizers are better than those obtained using the three-stages S-MSVQ quantizer (Fig. 4.15). *Good* PESQ scores are obtained for all the studied S-MSVQ quantizers and C values. For two-stages S-MSVQ, the average PESQ scores are better than those obtained for three-stages S-MSVQ. As the best PESQ value is achieved using two-stages three-splits S-MSVQ and $C = 20\%$, the latter parameters were selected and used in the CS-speech coding module integrated in improved version of the new end-to-end communication system. The average PESQ and SII scores are listed in Table 4.4. This table also highlights a comparison with the results obtained using the proposed communication systems with $BR = 8.85$ kbit/s and some other CS-based speech compression methods. The improved version of the proposed communication system design allows *good* quality of recovered speech, justified by a PESQ score close to 3.14, and *fair* intelligibility (SII= 0.47) for a $BR = 8.85$ kbit/s. Results in this table confirm that the improved version of the proposed communication system leads to an enhanced quality and intelligibility of the recovered speech compared to the, firstly, proposed system (PESQ= 2.41 and SII= 0.45). Additionally, the obtained average PESQ scores are better than those achieved using the speech compression methods presented in (Giacobello et al., 2012; Al-Azawi and Gaze, 2018; Ji, Zhu, and Champagne, 2019).

The method presented in (Giacobello et al., 2012) performs speech coding based on sparse linear prediction. The PESQ value obtained for a $BR = 10.1$ kbits/s is close to 4.13 ± 0.13 in noise-free conditions, and it is close to 3.21 ± 0.19 in the presence of AWGN for an SNR of 10 dB. The latter value is close to our results (PESQ=3.14) obtained for lower $BR = 8.85$ kbit/s and in more dramatic communication scenario considering multipath fading effects (in addition to AWGN).

In (Al-Azawi and Gaze, 2018), the authors proposed a method of simultaneously speech compression and encryption. They didn't use a quantization scheme. In addition, noisy communication channel was not considered in their tests. They obtained a PESQ score of 4.43 for a $C = 60\%$. For lower compression ratio $C = 40\%$, they achieved a PESQ value of 2.73. Although we used a quantization scheme

(which engenders a quantization error), lower $C = 20\%$, and we considered a 10-dB Rayleigh channel, our average PESQ value 3.14 is better.

Ji *et al.* in (Ji, Zhu, and Champagne, 2019), proposed a new dictionary leaning method, and presented a new compressive speech sensing. They also proposed a new linear predictive coding (LPC) coefficients extraction technique for voiced and unvoiced speech. Considering AWGN and for a $C = 30\%$, the obtained PESQ scores are 2.4 and 2.25 for males and females, respectively. Even with higher $C = 80\%$, they achieved PESQ values of 2.71 and 2.8 for males and females, respectively. In our case, we considered the multipath fading effect in addition to AWGN, and with lower $C = 20\%$ we obtained a better PESQ score of 3.14.

About Complexity and Security Aspects

A reduction in processing time (in the new communication system design) is achieved using the proposed CS-speech coding. Since CS uses advanced reconstruction procedures to compensate the lack of measurements, the computational burden is shifted from the transmitter to the receiver. Nevertheless, the overall load of the design is implicitly reduced as CS affects the entire mobile communication system. Indeed, the following three impacts influenced our communication system load, in addition to reducing the energy consumption, the memory requirements, and the hardware and software costs. Firstly, our communication system process less samples (i.e., faster acquisition and smaller traffic volume); hence less computational burden is carried. Secondly, with the simultaneous acquisition and compression in our design (i.e., proposed CS-source coding), one can avoid the utilization of costly-computational compression techniques (e.g., AMR-WB). Therefore, a separate module for compression is not needed (i.e., less complexity for the system). Thirdly, the secrecy properties of CS measurements permit joint compression and encryption (again at the acquisition stage), eliminating the additional computational cost of a separate encryption protocol.

Indeed, the randomness of the sensing matrix leads naturally to an implicit encryption (as evidence, see (Orsdemir et al., 2008; Cambareri et al., 2015; Bianchi, Bioglio, and Magli, 2014; Bianchi, Bioglio, and Magli, 2016) which analyse the security properties of CS. Some results for speech encryption application are reported in (Cambareri et al., 2015)). In fact, random matrices are generated by a pseudo-random number generator using an initial seed. The latter is shared between the transmitter and the receiver (shared secret). The compressed sensing measurements are transmitted, and its intended receivers may recover \mathbf{s} by knowing \mathbf{y} , and by having a prior agreement on the initial seed (or shared secret) that is necessary to

reproduce Φ (note here that sharing only the seed instead of the matrix Φ reduces the memory requirements which makes another advantage of this scheme). Hence the CS measurements \mathbf{y} can be considered as an encrypted representation of the original signal \mathbf{s} . In other words, the security of CS relies on the fact that an attacker has no knowledge about the sensing matrix Φ since it does not have the seed used to generate it. Consequently, CS is considered as a symmetric encryption scheme (i.e., private key cryptosystem), where \mathbf{s} is the plaintext, \mathbf{y} is the ciphertext, and the linear transformation operated by the sensing matrix Φ (the acquisition process) is the encryption algorithm (the initial seed is the encryption key). Since Gaussian random matrix is used in our design, we can say that our communication system grants some level of security promoting implicitly an encryption at zero-cost. More information about CS in security field can be found in (Zhang et al., 2016).

4.6 Conclusion

A novel end-to-end communication scheme that uses a method based on CS for speech coding is presented in this chapter. An improved version is also introduced for lower bitrates. The communication systems are investigated for worst case communication conditions (Rayleigh environments). The channel model and the techniques chosen to mitigate the channel effects (to provide secure and robust communications) are described. Results for speech communication are also provided.

To sum up, the study outcomes led to the following key findings:

- Compressed sensing performs the acquisition and compression simultaneously; hence reducing the processing load, network resources, memory, and the power consumption, which offers significant advantages over current methods such as AMR-WB that allows a *good* speech quality for small bit-rates but requires high battery consumption. Moreover, by reducing the data amount one can transmit greater volume at a time, which increases the transmission rate.
- A reduced complexity speech coding system is achieved using optimal SQ or S-MSVQ of CS measurements and binary encoding. This further reduces the overall complexity of the communication system.
- The randomness of the acquisition matrix serves as encryption, which allows secure communications without additional costs.
- The use of ℓ_1 minimization and wavelet thresholding (for sparsity) enhances the robustness against noise.

- Diversity techniques keep their performances when using CS-speech codec in the communication system:
 1. More receiving antennas corresponds to better BERs.
 2. Maximal ratio combining method gives better results in terms of BERs than SC and EGC.
- *Good* perceptual evaluation of speech quality scores, and *fair* coherence speech intelligibility index values are achieved with the proposed mobile communication system and its improved version. For example, for a $BR = 12.8$ kbit/s, an average PESQ score close to 3.5 is achieved considering the proposed communication system in a 10 dB Rayleigh environment. The improved version of the proposed communication system leads to an average PESQ score of 3.14 for a $BR = 8.85$ kbit/s in a 10 dB Rayleigh environment.
- The improved version of the proposed communication scheme allows enhancing the performance of the latter.
- In degraded communication environments and lower bitrates, the average PESQ values obtained using our design are better than those obtained using several recently proposed CS-speech coding methods. For example, for $C = 30\%$, a PESQ scores of 2.40 and 2.25 were obtained for male and female speech using the method proposed in (Ji, Zhu, and Champagne, 2019) in AWGN channel. In our case, we obtained a PESQ value of 3.14 for $C = 20\%$ in 10 dB Rayleigh channel.

In the following chapter, another application of compressed sensing in another area (speech enhancement) will be presented.

Chapter 5

Speech Enhancement Using Compressed Sensing

5.1 Introduction

In this chapter, a focus is made on the speech enhancement, which is an important task in many areas including speech coding and communication. The sparsity of speech (in a suitable sparse basis or an overcomplete dictionary) has been exploited to design new methods for speech enhancement based on compressed sensing (CS). These alternative methods have proven better performance than conventional speech enhancement methods that have been widely used such as, spectral subtraction, Wiener filtering, wavelet denoising, *etc.* Here, we propose a novel speech enhancement approach based on CS (Haneche, Boudraa, and Ouahabi, 2020). Details of the proposed CS-based speech enhancement method are presented after a brief description of some conventional techniques. Numerical results and discussion are also provided later in the chapter.

5.2 Motivation

Speech has been the primary form of communication between humans since the beginning of human civilization. Additive background noise, coming from environment such as in restaurants, market, and factory, can mask the speech signal. This effect reduces the quality, makes difficult listening, and gives poor performance in many applications. Hence, speech enhancement is required. The main objective of speech enhancement is to improve the perceptual aspects of speech such as quality, which reduces the listener fatigue (Hu and Loizou, 2008). Several areas of signal processing, including speech coding, automatic speech recognition, biomedical applications, hearing aid, mobile radio communications, teleconferencing systems,

military communications, *etc.*, are concerned by speech enhancement. There are two main categories of speech enhancement: multiple channel and single channel approaches (Benesty, Makino, and Chen, 2005). Multi-channel solutions require more than one microphone for observations. They exploit the spatial diversity to separate the signal of interest from noise. Single channel techniques offer a much more computationally efficient solution since only one microphone is used (Paliwal, Wójcicki, and Schwerin, 2010). These methods prevail, because of their low cost and simplicity. Popular single channel speech enhancement techniques are based on spectral subtraction (Boll, 1979; Paliwal, Wójcicki, and Schwerin, 2010), filtering (Almajai and Milner, 2011), wavelet shrinkage (Donoho, 1995; Ouahabi, 2012), or statistical modeling methods (Loizou, 2007):

- Spectral subtraction was firstly suggested by Boll (Boll, 1979), and has gained huge acceptance due to its simplicity (Paliwal, Wójcicki, and Schwerin, 2010). Basically, the enhanced signal can be obtained by subtracting the estimated spectral noise from the noisy spectral observations. An accurate noise spectral component estimation is needed. Typically, a voice activity detector (VAD) is used to detect the presence of speech and non-speech periods for the estimation of noise statistics.
- Filtering, such as Wiener filtering and Kalman filter-based methods, estimates clean speech signals by minimizing an estimation error analytically (Almajai and Milner, 2011). These methods are more effective to denoise stationary noises than non-stationary noises.
- Wavelet denoising was introduced by Donoho (Donoho, 1995). It is based on the assumption that low amplitude wavelet coefficients represent noise. The enhanced signal is obtained by transform back the wavelet coefficients to time domain after performing soft/hard thresholding.
- Statistical modeling methods have an improved denoising capacities for non-stationary noises by employing statistical models to maximize certain statistical criteria, e.g., maximum likelihood (ML) (Loizou, 2007).

An alternative to these conventional methods is a technique based on CS. The idea is to exploit the sparsity of speech in the chosen dictionaries. Few important CS-based speech enhancement methods have been proposed (Sigg, Dikk, and Buhmann, 2012; Jančovič, Zou, and Köküer, 2012; Low, Pham, and Venkatesh, 2013; Wu, Zhu, and Swamy, 2013; Wu, Zhu, and Swamy, 2014; Ramdas, Mishra, and Gorthi, 2015; Luo et al., 2016; Wang et al., 2016). They seek to estimate a sparse representation of the clean speech by reconstructing only the sparse components (speech) from the mixture

(speech and noise). In fact, CS provides an efficient transform to the measurement domain, while the recovery algorithm provides a way to return to the data domain. Nevertheless, when the input SNR is low, the energy of the noise is relatively large, and this results in degraded reconstruction quality (Wang, Wang, and Hao, 2019). Hence, the performances of speech enhancement by sparse recovery are reduced for low SNR scenarios (Wang et al., 2016). To address this issue, we aim to exploit the information about the noise, in the measurement domain, besides sparse recovery in a CS-based speech enhancement framework. The idea is to subtract the noise measurement before performing the CS reconstruction to reduce the impact of the noise; hence more accurate sparse recovery can be achieved.

5.3 Brief Description of Some Conventional Speech Enhancement Methods

Additive noise coming from environment introduces an annoying distortion in the speech signal. Therefore, speech should often be enhanced before further processing. Let \mathbf{s} be a speech signal affected by the additive noise \mathbf{n} giving the noisy speech \mathbf{x} that can be represented as

$$\mathbf{x} = \mathbf{s} + \mathbf{n} \quad (5.1)$$

Speech enhancement methods aim to extract \mathbf{s} from the mixture \mathbf{x} . In the subsequent subsections, some classical speech enhancement methods; namely spectral subtraction, Wiener filtering, and wavelet denoising are described.

5.3.1 Spectral Subtraction

Spectral subtraction is historically one of the oldest algorithms, its main properties are its easy implementation and minimal complexity. It is a single channel speech enhancement technique. In spectral subtraction the enhanced speech is synthesized from the spectrum obtained by subtracting the noise spectrum from the input noisy signal spectrum. Therefore, the magnitude of the noise spectrum has to be estimated based on the available noisy speech signal, and this is a challenging task. Important assumptions in this method are the stationarity of speech and the non-correlation of noise with the clean signal (Upadhyay and Jaiswal, 2016). Considering speech segments of 20 – 30 ms, the frequency representation is described as

$$\mathbf{X} = \mathbf{S} + \mathbf{N} \quad (5.2)$$

where \mathbf{X} , \mathbf{S} , and \mathbf{N} are the spectrograms of signals in equation (5.1). The spectrogram of the enhanced speech is defined as

$$\hat{\mathbf{S}} = |\hat{\mathbf{S}}| \exp(j\varphi_x) \quad (5.3)$$

where φ_x is the phase spectra of the noisy speech \mathbf{x} , and $|\hat{\mathbf{S}}|$ is the estimated magnitude spectrogram obtained by

$$|\hat{\mathbf{S}}| = |\mathbf{X}| - |\hat{\mathbf{N}}| \quad (5.4)$$

where $|\hat{\mathbf{N}}|$ is the magnitude of the estimated noise spectrogram. The noise spectrum is estimated during pauses. Voice activity detection is used for detecting the presence/absence of speech. The time representation of the enhanced speech is obtained by inverse Fourier transformation of $\hat{\mathbf{S}}$.

Spectral subtraction introduces a new type of noise commonly referred to in the literature as “musical noise”. Another minor shortcoming of the spectral subtraction algorithm is the use of noisy phase that produces a roughness in the quality of the synthesized speech. As the presence of noise in the phase does not contribute much to the degradation of speech quality, and because the clean phase Estimation greatly increases the complexity of the enhancement algorithm, much effort has been focused on finding methods to reduce musical noise. Berouti and Boll’s subtraction methods (Berouti, Schwartz, and Makhoul, 1979; Boll, 1979) form the basis of spectral subtraction for the new methods. For example, Paliwal in 2010 (Paliwal, Wójcicki, and Schwerin, 2010) proposed to apply spectral subtraction algorithm in the modulation domain. Kamath and Loizou suggested to divide the spectrum into frequency bands and then to apply different nonlinear rules in each band (Kamath and Loizou, 2002). Yet, Virag suggested the use of a psychoacoustical model so as to render the residual noise inaudible (Virag, 1999).

5.3.2 Wiener Filtering

The Wiener filter is an optimal filter that minimize the MSE criterion. It differs from spectral subtraction by the fact that it gives the minimum MSE estimate of the short-time Fourier transform (STFT) whereas the spectral subtraction obtains the minimum MSE estimate of the short-time spectral magnitude without changing the phase (Vaseghi, 2005). The gain function of Wiener filter can be defined as (Upadhyay and Jaiswal, 2016)

$$\mathbf{G} = \frac{\mathbf{P}_s}{\mathbf{P}_x} = \frac{\mathbf{P}_x - \mathbf{P}_n}{\mathbf{P}_x} \quad (5.5)$$

here \mathbf{P}_s , \mathbf{P}_n , and \mathbf{P}_x are the power spectral density of the clean speech, the noise, and the noisy speech, respectively. The enhanced speech is estimated as (Upadhyay and Jaiswal, 2016)

$$\hat{\mathbf{P}}_s = \mathbf{G}\mathbf{P}_x \quad (5.6)$$

$$|\hat{\mathbf{s}}| = TF^{-1}(\sqrt{\hat{\mathbf{P}}_s}) \quad (5.7)$$

Taking into account the (noisy) phase, we obtain

$$\hat{\mathbf{s}} = |\hat{\mathbf{s}}|e^{j\varphi_x} \quad (5.8)$$

5.3.3 Wavelet Denoising

Wavelet shrinkage denoising is considered as a non-parametric method (Ahmed et al., 2015; Ouahabi, 2013). It offers an adaptive resolution better than STFT (Ouahabi, 2012). Speech signal has sparse representation in wavelet basis. Thus, one can eliminate low coefficients, which carry non significant information (mostly represent noise) to have a enhanced speech. Wavelet-based denoising method consists of three steps:

1. Perform forward wavelet transform on noisy speech (Eq. (5.1)), to obtain

$$\mathbf{W}_x = \mathbf{W}_s + \mathbf{W}_n \quad (5.9)$$

where \mathbf{W}_x , \mathbf{W}_s , and \mathbf{W}_n denote the wavelet transform representations of noisy, clean, and noise signals, respectively.

2. Apply a threshold γ , using a thresholding function $T(\cdot)$, on the wavelet coefficients $T(\mathbf{W}_x, \gamma)$. Several threshold estimation methods exist, e.g., universal, visushrink, minimax, and heuristic sure thresholds. The method proposed by Donoho and Johnstone (Donoho and Johnstone, 1994; Donoho, 1995), called universal threshold, is widely used. It is defined by Eq. (4.3). This threshold is used in a thresholding function that can be hard, soft, or semi-soft thresholding function. Hard thresholding is defined as

$$T_{hard}(\mathbf{W}_x, \gamma) = \begin{cases} \mathbf{W}_x & \text{if } |\mathbf{W}_x| > \gamma \\ 0 & \text{otherwise} \end{cases} \quad (5.10)$$

where $T_{hard}(\cdot)$ denotes the hard thresholding function and γ the universal threshold.

The estimated wavelet representation of the enhanced speech is described as

$$\widehat{\mathbf{W}}_s = T_{hard}(\mathbf{W}_x \gamma) \quad (5.11)$$

3. Perform inverse wavelet transform to $\widehat{\mathbf{W}}_s$.

5.4 Compressed Sensing-Based Speech Enhancement

Compressed sensing can be used for speech enhancement thanks to sparsity of clean speech in a chosen basis. The basic CS-based speech enhancement steps are:

1. Sparse representation: By projecting the noisy speech \mathbf{x} in a suitable sparse basis Ψ (e.g, wavelet, time-frequency, ... representation) as defined by Eq. (2.2), and performing hard thresholding (using the universal threshold) to eliminate non-significant coefficients, and hence reinforce sparsity.
2. Acquisition: By multiplying the sparse vector of coefficients θ with a random Gaussian matrix Φ (see Eq. (2.18)).
3. Sparse recovery: By applying a CS reconstruction algorithm (refer to Section 2.3.3 for more details), for example ℓ_1 minimization defined in Eq. (2.24).

The enhanced speech $\widehat{\mathbf{s}}$ is obtained by inverse sparse transform (e.g, inverse wavelet transform, inverse short-time Fourier transform, ...) of the recovered sparse vector $\widehat{\theta}$.

In the following section, we present our proposed CS-based speech enhancement method.

5.5 Proposed compressed sensing-based speech enhancement method

Sparsity plays an important role in CS. It is the first condition that must be verified to allow CS reconstruction. The fact that most real-life signals are sparse in some transform bases has encouraged the usage of CS in many applications, including speech enhancement. Speech signals are sparse in the time-frequency domain (Low, Pham, and Venkatesh, 2013; Wang et al., 2016); hence our choice to perform the enhancement in the time-frequency domain. In this section, the general description as well as the mathematical formulation of the proposed CS-based speech enhancement approach are given.

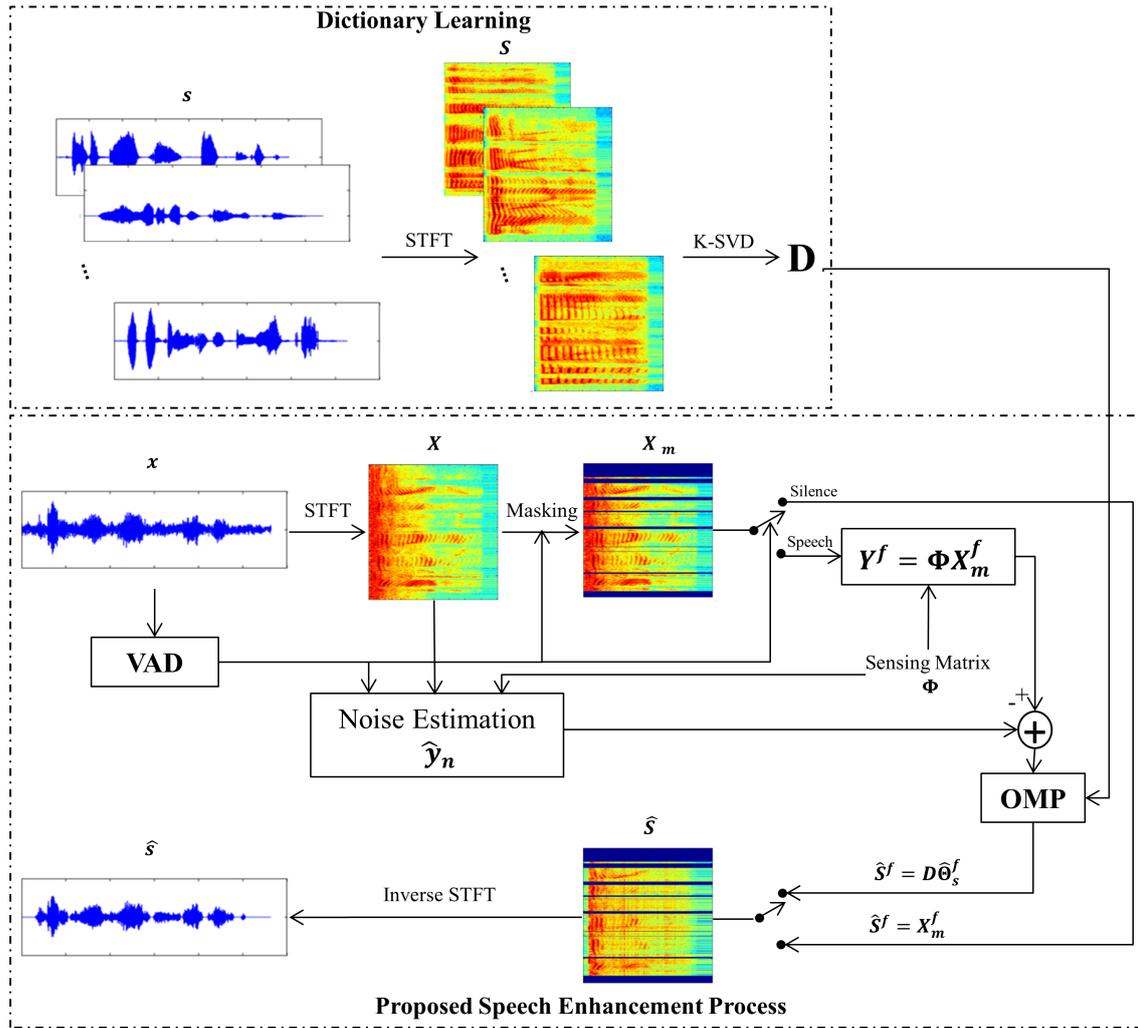


FIGURE 5.1: Block Diagram of the proposed CS-based speech enhancement method. Note that for each frame, a decision speech/silence is performed and a processing is done depending on the result.

5.5.1 System Overview

We propose a speech enhancement method based on CS by performing noise subtraction in the measurement domain. Fig. 5.1 presents our enhancement technique. This approach is divided into two stages: dictionary learning and speech enhancement procedure. For dictionary learning we used K-SVD algorithm to train, in the frequency domain, an overcomplete dictionary \mathbf{D} that would be used for CS recovery. This dictionary can sparsely represent the speech signals with a tolerable error. The proposed speech enhancement process is performed on the overlapped frames in the frequency domain. First, speech segmentation using overlapped Hamming windows is performed. For each frame f , we compute the frequency representation (in other words, STFT). The noise measurement can be performed based on noise-only frames. The frames corresponding to the first 100 – 300 ms of every noisy speech

are usually considered as silence; hence they contain only noise (Boll, 1979; Cole, Karam, and Aglan, 2008; Seyedtabaee and Goodarzi, 2010). In this work, the first 250 ms are assumed to be the initial silence, and the corresponding N_{IS} frames are used to compute the initial noise $\hat{\mathbf{y}}_n$ (in the measurement domain) by averaging the observation vectors of these segments. Next, VAD is performed, and a mask function \mathbf{M} is created depending on the VAD output. If the current frame is classified as speech ($\text{VAD}(f) = 1$), masking is performed on the speech spectrum, and the noise (in the measurement domain) is subtracted from the observation vector of the noisy speech. The enhanced speech spectrum is obtained by CS recovery using orthogonal matching pursuit (OMP). Otherwise, if the current frame is considered as silence ($\text{VAD}(f) = 0$), the noise measurement are updated, and the mask is applied on the spectrum. In this case, the enhanced speech spectrum is the masked spectrum. After that, we return to the time domain, and we overlap-add speech segments to obtain the enhanced speech (in other words, inverse STFT). Algorithm 9 details the speech enhancement procedure.

5.5.2 Mathematical Formulation of the Proposed Speech enhancement process

Assuming a noisy speech $\mathbf{x} \in \mathbb{R}^N$, it can be represented as defined in Eq. (5.1). Considering overlapped segments and passing to the frequency domain, we obtain the spectrogram $\mathbf{X} \in \mathbb{R}^{B \times F}$ (see Eq. (5.2)), where B is the number of frequency bins, and F is the number of frames. \mathbf{X} can be spanned by few atoms from an overcomplete dictionary trained using clean speech spectrograms, giving hence the sparse representation Θ

$$\mathbf{X} = \mathbf{D}\Theta \quad (5.12)$$

where $\mathbf{D} \in \mathbb{R}^{B \times T}$ is the overcomplete dictionary, and T is the number of atoms. Θ is a $T \times F$ matrix containing the sparse coefficients vectors Θ^f . Note that

$$\forall f = 1, \dots, F \quad \|\Theta^f\|_0 \leq K$$

where Θ^f is the f th frame of Θ , and K is the sparsity.

The CS acquisition (sampling process) holds the matrix of observation vectors \mathbf{Y}_x defined as

$$\mathbf{Y}_x = \Phi\mathbf{X} = \Phi\mathbf{D}\Theta \quad (5.13)$$

where $\Phi \in \mathbb{R}^{M \times B}$ is random Gaussian sensing matrix ($M < B$), and \mathbf{Y}_x is the matrix of dimension $M \times F$ containing the observation vectors (Measurement vectors). The

Algorithm 9 Proposed CS-based speech enhancement procedure**Input:** \mathbf{x} the noisy speech and \mathbf{D} the overcomplete dictionary**Output:** The enhanced speech $\hat{\mathbf{s}}$

```

1: Speech segmentation ▷ Hamming window
2: Create a sensing matrix  $\Phi$  ▷ random Gaussian matrix
3: Perform noise measurement  $\hat{\mathbf{y}}_n$  ▷ for the 1st  $N_{IS}$  frames (Eq. (5.15))
4: for each frame  $f$  do
5:   Compute spectrum  $\mathbf{X}^f$ 
6:   Perform VAD
7:   Create a mask  $\mathbf{M}^f$  ▷ with regard to Eq. (5.18)
8:   if VAD( $f$ ) = 1 then ▷ speech frame detected
9:     Apply the mask to the spectrum as
        $\mathbf{X}_m^f \leftarrow \mathbf{M}^f \mathbf{X}^f$  ▷ element-wise multiplication
10:    Estimate the clean speech frame measurement  $\hat{\mathbf{y}}_s$  as
        $\hat{\mathbf{y}}_s \leftarrow \Phi \mathbf{X}_m^f - \hat{\mathbf{y}}_n$  ▷ noise subtraction in the measurement domain
11:    Perform OMP to obtain the the sparse coefficients vector  $\hat{\Theta}_s^f$ 
12:    Compute the enhanced spectrum  $\hat{\mathbf{S}}^f$  as
        $\hat{\mathbf{S}}^f \leftarrow \mathbf{D} \hat{\Theta}_s^f$ 
13:   else ▷ silence frame detected
14:     Update the noise measurement  $\hat{\mathbf{y}}_n$  ▷ with regard to Eq. (5.16)
15:     Apply the mask to the spectrum as
        $\mathbf{X}_m^f \leftarrow \mathbf{M}^f \mathbf{X}^f$  ▷ element-wise multiplication
16:     Compute the enhanced spectrum  $\hat{\mathbf{S}}^f$  as
        $\hat{\mathbf{S}}^f \leftarrow \mathbf{X}_m^f$  ▷ the enhanced spectrum is the masked spectrum
17:   end if
18: end for
19: Return to the time domain
20: Overlap-add the enhanced speech segments
21: return  $\hat{\mathbf{s}}$ 

```

speech spectra are independently projected on the measurement domain using the same sensing matrix Φ , *i.e.*, for each f th frame \mathbf{X}^f from \mathbf{X} corresponds an observation vector $\mathbf{Y}_x^f \in \mathbb{R}^{M \times 1}$ which is the f th column of \mathbf{Y}_x .

By substituting \mathbf{X} (see Eq. 5.2), the latter equation is equivalent to

$$\mathbf{Y}_x = \Phi \mathbf{S} + \Phi \mathbf{N} = \mathbf{Y}_s + \mathbf{Y}_n \quad (5.14)$$

where \mathbf{Y}_s and \mathbf{Y}_n are the matrices containing the observation vectors corresponding to the clean speech and the noise, respectively. In other words, they are the projections of \mathbf{S} and \mathbf{N} in the measurement domain.

To perform noise subtraction in the measurement domain, we need to estimate the observation vector of the noise (*i.e.*, estimate the noise in the measurement domain).

This can be done during pauses. We used VAD technique to detect voice/silence frames.

The principle of VAD devices is based on features extraction and comparison with thresholds computed from noise-only periods. Generally, VAD has a binary output. Voice activity is detected (VAD output = 1) if the features values exceed the thresholds. Otherwise, silence is detected (VAD output = 0). More accurate thresholds provide more reliable voice detection results. After many experiments, we chose the algorithm proposed by Verteletskaya (Verteletskaya and Sakhnov, 2010) for voice/silence detection. This algorithm is based on periodicity detection, as well as the high-frequency to low-frequency signal energy ratio and full-band energy computation.

The noise in the measurement domain was estimated by frame-averaging the observation vectors corresponding to the silence segments. Initially, we considered the first N_{IS} frames as silence, and we estimated noise measurements as

$$\hat{\mathbf{y}}_n(i) = \frac{1}{N_{IS}} \sum_{f=1}^{N_{IS}} \mathbf{Y}_x(i, f), \quad i = 1, \dots, M \quad (5.15)$$

where f is the frame index, M is the number of measurements, and \mathbf{Y}_x is the matrix of observation vectors corresponding to the noisy speech \mathbf{x} . Usually, it is assumed that the first 100 – 300 ms of the speech signal do not contain speech (Boll, 1979; Cole, Karam, and Aglan, 2008; Seyedtabaee and Goodarzi, 2010). Therefore, in this work we considered that the number of the initial silence frames N_{IS} corresponds to the initial 250 ms, which is a noise-only signal. $\hat{\mathbf{y}}_n$ is updated when a silence frame f is detected ($\text{VAD}(f) = 0$). Hence, the frame-averaged noise observation vector is expressed as

$$\hat{\mathbf{y}}_n(i) = \frac{1}{N_S} [(N_S - 1)\hat{\mathbf{y}}_n(i) + \mathbf{Y}_x(i, f)] \quad (5.16)$$

where N_S is the total number of silence segments.

After estimating noise in the measurement domain, we applied a mask function \mathbf{M} to the spectrogram \mathbf{X} . This function returns to zero all the silence frames giving, hence, \mathbf{X}_m defined as

$$\mathbf{X}_m = \mathbf{M}\mathbf{X} \quad (5.17)$$

with

$$\mathbf{M}(b, f) = \begin{cases} 1 & \text{if } \text{VAD}(f) = 1 \\ 0 & \text{if } \text{VAD}(f) = 0 \end{cases} \quad (5.18)$$

where $b = 1, \dots, B$ is the frequency bins index, and $f = 1, \dots, F$ is the frame index. The multiplication is performed in an element-wise manner.

The noise subtraction (in the measurement domain) was performed for each active-speech frame as follows

$$\widehat{\mathbf{Y}}_s^f = \mathbf{Y}^f - \widehat{\mathbf{y}}_n \quad \text{if } \text{VAD}(f) = 1 \quad (5.19)$$

where f is the frame index, and $\widehat{\mathbf{Y}}_s^f$ and \mathbf{Y}^f are the f th frames of $\widehat{\mathbf{Y}}_s$ and \mathbf{Y} , respectively. $\widehat{\mathbf{Y}}_s$ and \mathbf{Y} denote the matrices containing the observation vectors corresponding to the enhanced and masked spectrograms, respectively. $\mathbf{Y} = \Phi \mathbf{X}_m$.

After performing noise subtraction in the measurement domain, OMP algorithm was applied to recover the sparse coefficients vectors. OMP is a greedy algorithm that is proved to be an attractive alternative of basis pursuit (Tropp and Gilbert, 2007). It is faster and easy to implement. The problem can be formulated as

$$\widehat{\Theta}_s^f = \arg \min_{\Theta_s^f} \|\widehat{\mathbf{Y}}_s^f - \Phi \mathbf{D} \Theta_s^f\|_2 \quad \text{s.t.} \quad \|\Theta_s^f\|_0 \leq K \quad (5.20)$$

or as

$$\widehat{\Theta}_s^f = \arg \min_{\Theta_s^f} \|\Theta_s^f\|_0 \quad \text{s.t.} \quad \|\widehat{\mathbf{Y}}_s^f - \Phi \mathbf{D} \Theta_s^f\|_2 \leq \sigma \quad (5.21)$$

where $\widehat{\Theta}_s^f$ is the f th frame of the estimated sparse coefficients matrix $\widehat{\Theta}_s$.

The enhanced spectrogram is described as

$$\widehat{\mathbf{S}}^f = \begin{cases} \mathbf{D} \widehat{\Theta}_s^f & \text{if } \text{VAD}(f) = 1 \\ \mathbf{X}_m^f & \text{if } \text{VAD}(f) = 0 \end{cases} \quad (5.22)$$

where $\widehat{\mathbf{S}}^f$ and \mathbf{X}_m^f are the f th frames of $\widehat{\mathbf{S}}$ and \mathbf{X}_m respectively.

The enhanced speech $\widehat{\mathbf{s}}$ was obtained by performing the inverse STFT.

Experiments have been done to assess the proposed CS-based speech enhancement method. The simulation setup, performance measures, and the results and discussion are presented in the following section.

5.6 Experiments

The performances of our approach are evaluated in comparison with recent state-of-the-art methods, using various interferer signals in terms of three objective measures,

namely, segmental SNR, perceptual evaluation of speech quality, and short-time objective intelligibility.

5.6.1 Experimental Setting

We used clean speech from NOIZEUS database (Hu and Loizou, 2007). The sampling frequency of these signals is 8 kHz. The speech signals were partitioned into two disjoint groups: a training set and a test set. Ten utterances (five females and five males) were used for training. The data in the test set were from 20 speakers (10 females and 10 males). The test clean data were corrupted by adding various noises with various SNR levels to simulate adverse environments. These noises include *babble*, *police siren*, *piano*, *market*, *factory*, and *white Gaussian noise*. *Babble*, *factory*, and *white Gaussian noise* were obtained from NOISEX-92 database (Varga and Steeneken, 1993), and *police siren*, *piano* and *market* noises from (Freesound 2019). All the noise data were downsampled to 8 kHz. Recall that the initial 250 ms (corresponding to N_{IS} frames) of each noisy speech signal were assumed to be a noise-only signal, and used to perform the initial noise measurement.

The overcomplete dictionary was learned using spectrograms of clean training set. 512 frequency bins were considered for the STFT with 32 ms Hamming window and 4 ms shift. K-SVD algorithm was adopted using 1024 atoms, because it offers trade-off between the performance of the dictionary and training time. The number of iterations was set to 100 as in (Wang et al., 2016). To chose the sparsity value of K-SVD, many experiments were carried out. In these experiments, the overcomplete dictionaries were trained using different sparsity values of K-SVD, ranging from 10 to 50. The PESQ scores of the enhanced speech were computed for 10 dB *babble* noise (similar to (Wang et al., 2016)). They are presented in Table 5.1. The highest PESQ score is shown in bold. As evidenced in this table, the best PESQ score is obtained

TABLE 5.1: PESQ performances of the proposed CS-based speech enhancement method for different sparsity values of K-SVD in 10 dB *babble* noise.

Sparsity	10	15	20	25	30	35	40	45	50
PESQ	2.45	2.47	2.49	2.48	2.46	2.44	2.46	2.44	2.46

using the sparsity value of 20. This finding is in accordance with that reported in (Wang et al., 2016). Hence, the sparsity value of K-SVD was set to 20 in this work.

In the enhancement process, all the noisy utterances were transformed into the time-frequency domain using the STFT. A mask function \mathbf{M} was created based on

the VAD output, as defined by Eq. (5.18). The clean spectra were estimated as described in subsection 5.5.2. This method is based on CS and noise subtraction in the measurement domain. The noise measurements were initially estimated by averaging the observation vectors of the first N_{IS} silence segments, and then updated according to Eq. (5.16). VAD was used to detect the silence/speech segments. A random Gaussian matrix was used as sensing matrix. The number of measurement M was chosen (as in (Low, Pham, and Venkatesh, 2013)) to be 90% of the segment length (because the first purpose is enhancing speech rather than compressing it). OMP algorithm was used as a reconstruction algorithm. Inverse STFT was applied to obtain the enhanced speech in the time domain.

5.6.2 Performance Measures

Three objective measures are utilized to evaluate the performance of the proposed CS-based speech enhancement algorithm: the segmental SNR (SNRseg), the perceptual evaluation of speech quality (PESQ), and the short-time objective intelligibility (STOI). The segmental SNR is defined as the mean of SNRs of all the segments (frames). It is defined as (Loizou, 2007)

$$SNR_{seg} = \frac{10}{F} \sum_{f=0}^{F-1} \log_{10} \frac{\sum_{i=N.f}^{N.f+N-1} \mathbf{s}^2(i)}{\sum_{i=N.f}^{N.f+N-1} (\mathbf{s}(i) - \widehat{\mathbf{s}}(i))^2} \quad (5.23)$$

where \mathbf{s} and $\widehat{\mathbf{s}}$ are the clean and the enhanced signals, N is the frame length, and F is the number of frames.

The STOI (Taal et al., 2011) is an objective measure of speech intelligibility. It is highly correlated to human speech intelligibility score. The STOI score ranges from 0 (bad) to 1 (excellent). Fig. 5.2 gives more details on the STOI scale.

For more details about the PESQ scale, refer to Fig. 3.10.



FIGURE 5.2: The STOI scale.

5.6.3 Results and Discussion

We reduced the impact of the noise on the CS recovery in two main steps: (i) estimating the noise (in the measurement domain) during pauses, and subtracting it from the observation vectors of active speech frames before CS recovery;

(ii) applying a mask function that considers all the frequency bins of silence frames unreliable.

One can note from the aforementioned contributions that different processes are performed for speech and silence frames; hence the need to a VAD technique in the proposed method.

To illustrate the effect of the VAD, an experiment was carried out without the use of VAD. In other words, pure CS was used to enhance the noisy speech.

The performances are shown in Table 5.2. From these results, one can note that better performances, in terms of SNRseg improvements and PESQ, are achieved using the proposed CS-based speech enhancement method. Moreover, it is noticed that this improvement is more significant in low SNRs than in higher SNRs.

TABLE 5.2: SNRseg improvement, PESQ, and STOI performances of the CS-based speech enhancement method (without VAD), the proposed CS-based speech enhancement method, and the proposed CS-based speech enhancement method with the Oracle mask in *babble* noise.

Method		Noise level (dB)						
		-10	-5	0	5	10	15	20
SNRseg(dB)	CS (no VAD)	1.37	1.76	1.86	1.78	2.65	0.89	-1.5
	Proposed	3.87	3.75	3.64	3.40	3.01	1.48	-0.95
	Proposed (Oracle)	3.93	3.90	3.8	3.5	3.06	1.6	-0.94
PESQ	CS (no VAD)	0.83	1.13	1.37	1.65	2.27	2.47	2.61
	Proposed	1.21	1.54	1.89	2.21	2.49	2.64	2.80
	Proposed (Oracle)	1.26	1.6	1.96	2.28	2.52	2.69	2.85
STOI	CS (no VAD)	0.47	0.55	0.64	0.72	0.77	0.80	0.83
	Proposed	0.48	0.55	0.65	0.72	0.78	0.80	0.83
	Proposed (Oracle)	0.49	0.57	0.66	0.72	0.79	0.81	0.83

Nonetheless, VAD techniques may omit errors when classifying speech/silence segments, as in some cases they confuse between the noise and unvoiced sounds. Thus, a study on how sensitive is our method to the VAD errors is needed. To this end, a new experiment was performed using an ideal binary mask instead of the VAD-based mask within the proposed method. This ideal mask was created based on the Oracle estimator (i.e., assuming that the clean signal is known). Therefore, this approach can assert that our method achieves the best performances.

To explain further, Oracle estimators are algorithms that determine the separation parameters providing the best possible performance. The resulting estimates are called oracle estimates. In view of this definition, oracle estimators can be used, in an

evaluation context, only if reference source signals are available. The binary mask is a signal processing technique that retains the speech-dominated time-frequency regions, while zeroing out the noise-dominated time-frequency regions of noisy speech. The ideal or oracle binary mask is the best possible binary mask for a given signal. It can be computed only if the clean speech is known. In this (ideal) case, the oracle estimator makes the retain/discard decision based on the oracle (exact) SNR of each time-frequency bin, compared to a threshold called the local SNR criterion (LC). The ideal binary mask can be defined as

$$\mathbf{M}_{IBM}(b, f) = \begin{cases} 1 & \text{if } \text{SNR}(b, f) \geq LC \\ 0 & \text{otherwise} \end{cases} \quad (5.24)$$

where $b = 1, \dots, B$ is the frequency bins index, and $f = 1, \dots, F$ is the frame index. LC is the local SNR criterion in dB, and its value is often arbitrarily determined, since it has been believed that the specific choices of LC won't significantly affect the intelligibility of the processed speech, as long as the chosen LC value is not too low or too high. In other words, the speech intelligibility is largely independent of LC for a wide range of LC values (typically between -12 and 0 dB) (Brungart et al., 2006; Kjems et al., 2009). Other works found that the LC value for the ideal binary mask should be set to 5 dB below the overall SNR (Kjems et al., 2009; Wang, Narayanan, and Wang, 2014; Zhang and Shen, 2019). In our simulations, we set the value of LC to -5 dB. The ideal binary mask cannot occur in real life, but for several reasons this result is important. Particularly, it can be a criterion to estimate the performance of a practical algorithm (speech enhancement in our case), i.e., an upper bound.

The performance of the enhanced speech using the proposed CS-based speech enhancement method with the Oracle mask in the case of *babble* noise is evaluated in Table 5.2. The results indicate that the proposed method with the Oracle mask achieves the best performances, but overall, they are quite close to those of the proposed CS-based speech enhancement approach (with the VAD-based mask). This shows that the proposed method (using the chosen VAD technique) works successfully, i.e., it is robust to the VAD errors.

To evaluate the performance of the proposed CS-based speech enhancement method, we compared the results against several recent CS-based speech enhancement algorithms from the literature (see Table 5.3), including Wang16 (Wang et al., 2016), compressive sampling matching pursuit speech enhancement (CoSaMPSE) (Wu, Zhu, and Swamy, 2013; Wu, Zhu, and Swamy, 2014), generative dictionary learning (GDL) (Sigg, Dikk, and Buhmann, 2012), complementary joint sparse representations

TABLE 5.3: Methods used for comparison

Method	Based or Not on CS	Year	Reference
Wang16	Yes	2016	(Wang et al., 2016)
Compressive sampling matching pursuit speech enhancement (CoSaMPSE)	Yes	2013	(Wu, Zhu, and Swamy, 2013; Wu, Zhu, and Swamy, 2014)
Generative dictionary learning (GDL)	Yes	2012	(Sigg, Dikk, and Buhmann, 2012)
Complementary joint sparse representations (CJSR)	Yes	2016	(Luo et al., 2016)
Low13	Yes	2013	(Low, Pham, and Venkatesh, 2013)
Regularized non-negative matrix factorization with Gaussian mixtures and weighted Wiener filtering (RNG-WWF)	No	2017	(Chung, Plourde, and Champagne, 2017)
Supervised non-negative matrix factorization (supervised NMF)	No	2013	(Mohammadiha, Smaragdis, and Leijon, 2013)

(CJSR) (Luo et al., 2016), and the method proposed by Low (Low13) (Low, Pham, and Venkatesh, 2013). Other recent speech enhancement methods (not based on CS), namely, regularized non-negative matrix factorization with Gaussian mixtures and weighted Wiener filtering (RNG-WWF) (Chung, Plourde, and Champagne, 2017), and supervised non-negative matrix factorization (supervised NMF) (Mohammadiha, Smaragdis, and Leijon, 2013), were also considered for comparison. We chose these methods because they have been shown that they outperform the classical speech enhancement methods as well as many other speech enhancement techniques (Chung, Plourde, and Champagne, 2017; Wang et al., 2016; Wu, Zhu, and Swamy, 2013; Wu, Zhu, and Swamy, 2014; Low, Pham, and Venkatesh, 2013), e.g., regularized NMF (Grais and Erdogan, 2013), the maximum a posteriori estimator of the magnitude squared spectrum (Lu and Loizou, 2011), Wu11 (Wu, Zhu, and Swamy, 2011), the Wiener based method (Benesty, Makino, and Chen, 2005), etc.

Figs. 5.3, 5.4, and 5.5 compare, respectively, the segmental SNR improvement, PESQ, and the STOI values of the proposed algorithm with the CS-based speech enhancement methods proposed in (Wang et al., 2016; Wu, Zhu, and Swamy, 2014) for different noise types (*babble*, *police siren*, *piano*, and *market*) and various SNR levels ranging from -10 dB to 20 dB. For most noise scenarios, the SNRseg improvements

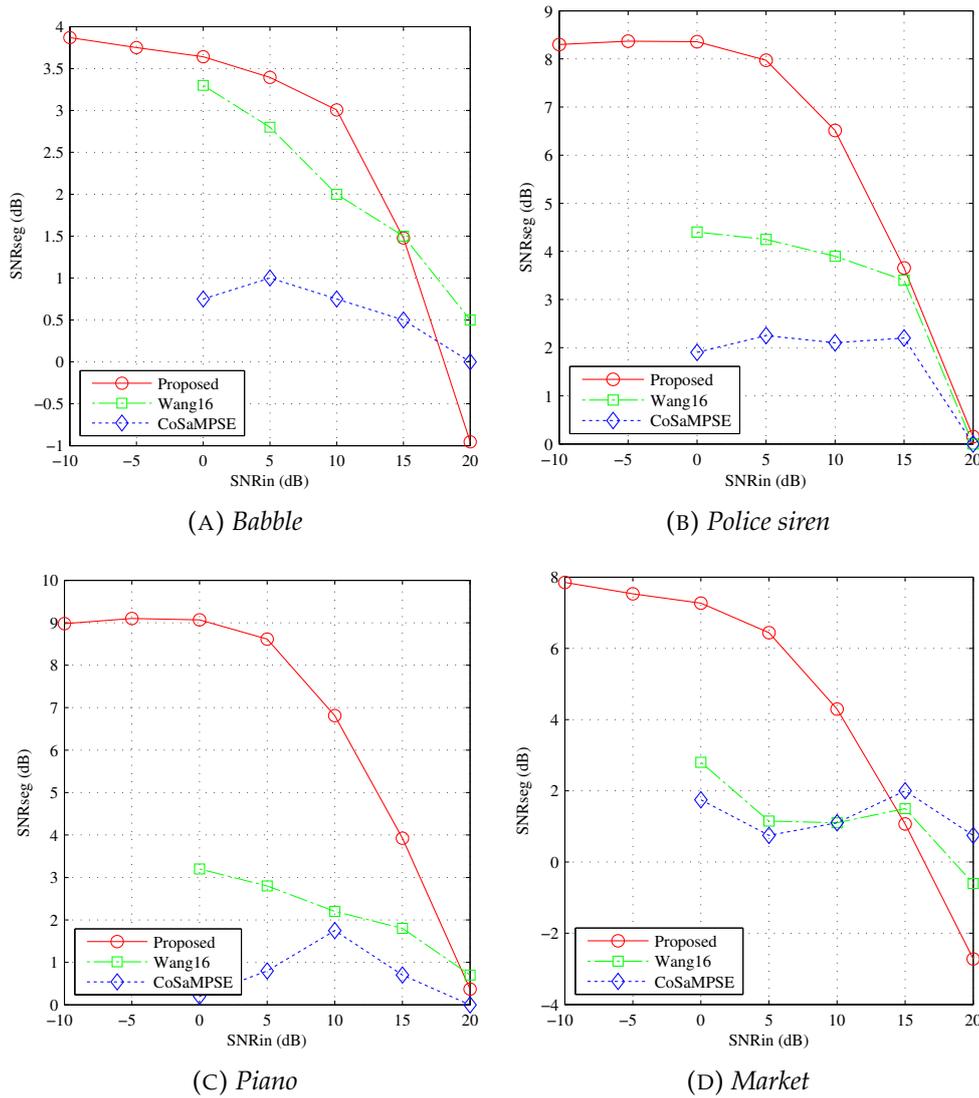


FIGURE 5.3: Segmental SNR improvements of the proposed CS-based speech enhancement method, Wang16 and CoSaMPSE in various noise conditions : (a) *babble*, (b) *police siren*, (c) *piano*, and (d) *market*. The segmental SNR improvements of Wang16 and CoSaMPSE are not available for the input SNRs -10 dB and -5 dB.

of the proposed method decrease at low rate for low input SNRs. For *babble* and *market* noises (Fig. 5.3 (a) and (d)), our SNRseg improvements are higher (significantly higher for *market* noise) than those of Wang16 and CoSaMPSE at various input SNRs levels (0–15 dB). For *police siren* and *piano* noises (Fig. 5.3 (b) and (c)), our SNRseg improvements are significantly higher than those of Wang16 and CoSaMPSE. From Fig. 5.4, we notice that for *babble* noise, *fair* PESQ scores ($1.5 \leq \text{PESQ} \leq 2.5$) are obtained for $-5 \leq \text{SNRs} \leq 10$ dB, and *good* PESQ scores ($2.5 \leq \text{PESQ} \leq 3.5$) are obtained for higher SNRs. For *police siren* noise, *fair* PESQ scores are obtained for $-8 \leq \text{SNRs} \leq 10$ dB, and *good* PESQ scores are obtained for higher SNRs. For the

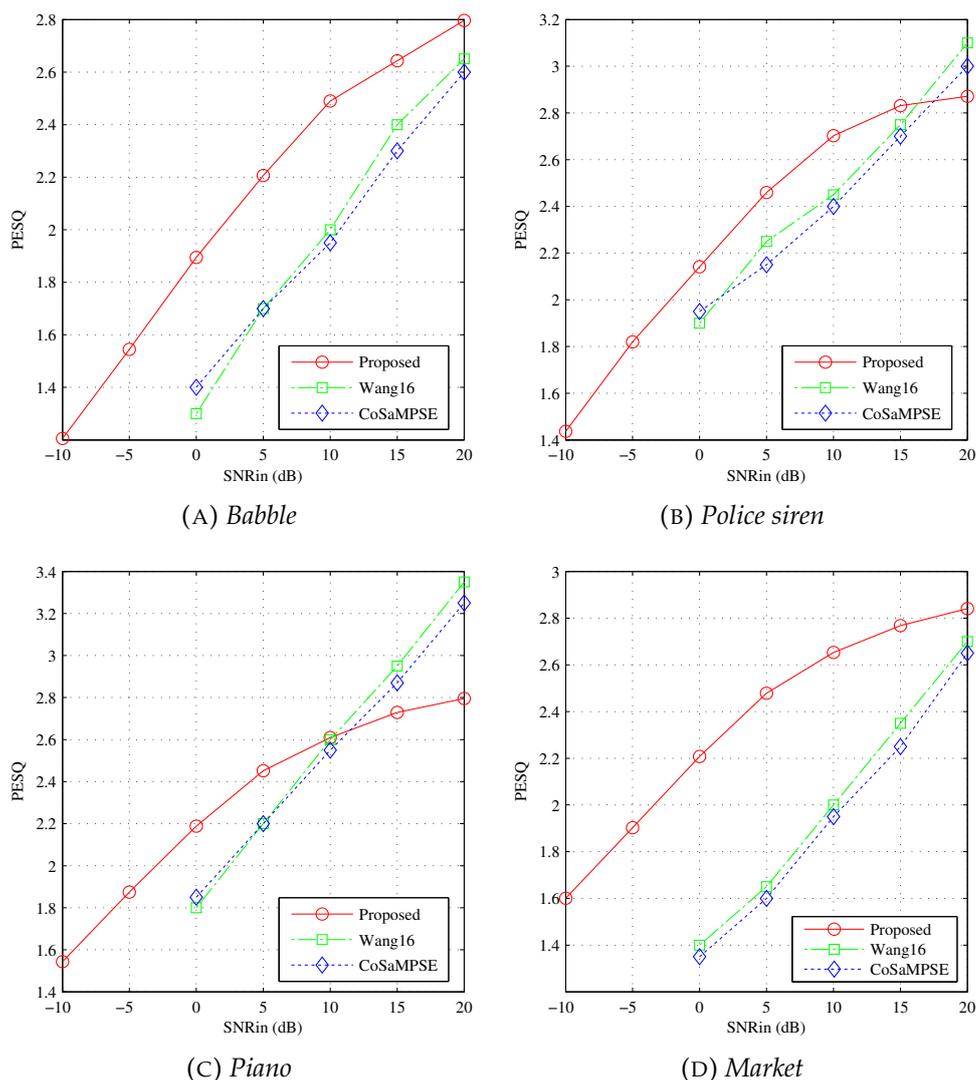


FIGURE 5.4: PESQ performances of the proposed CS-based speech enhancement method, Wang16 and CoSaMPSE in various noise conditions: (a) *babble*, (b) *police siren*, (c) *piano*, and (d) *market*. The PESQ scores of Wang16 and CoSaMPSE are not available for the input SNRs -10 dB and -5 dB.

other types of noises (*piano*, and *market* noises), *fair* PESQ values are obtained for low SNRs (≤ 5 dB), and *good* PESQ scores for higher SNR levels. Our CS-based speech enhancement method significantly outperforms Wang16 and CoSaMPSE techniques in *babble* and *market* noises (Fig. 5.4 (a) and (d)). It also provides better PESQ results for input SNR levels ≤ 15 dB with *police siren* noise, and ≤ 10 dB with the *piano* noise. In all noise scenarios, *good* STOI ($0.6 \leq \text{STOI} \leq 0.75$) and *excellent* STOI ($0.75 \leq \text{STOI} \leq 1$) values are obtained (Fig. 5.5). With *babble* noise (Fig. 5.5 (a)), they are better than the results of CoSaMPSE for SNR levels ≤ 10 dB, and comparable to the results of Wang16 for low input SNRs approaching 0 dB. For *piano* noise (Fig. 5.5 (c)), the

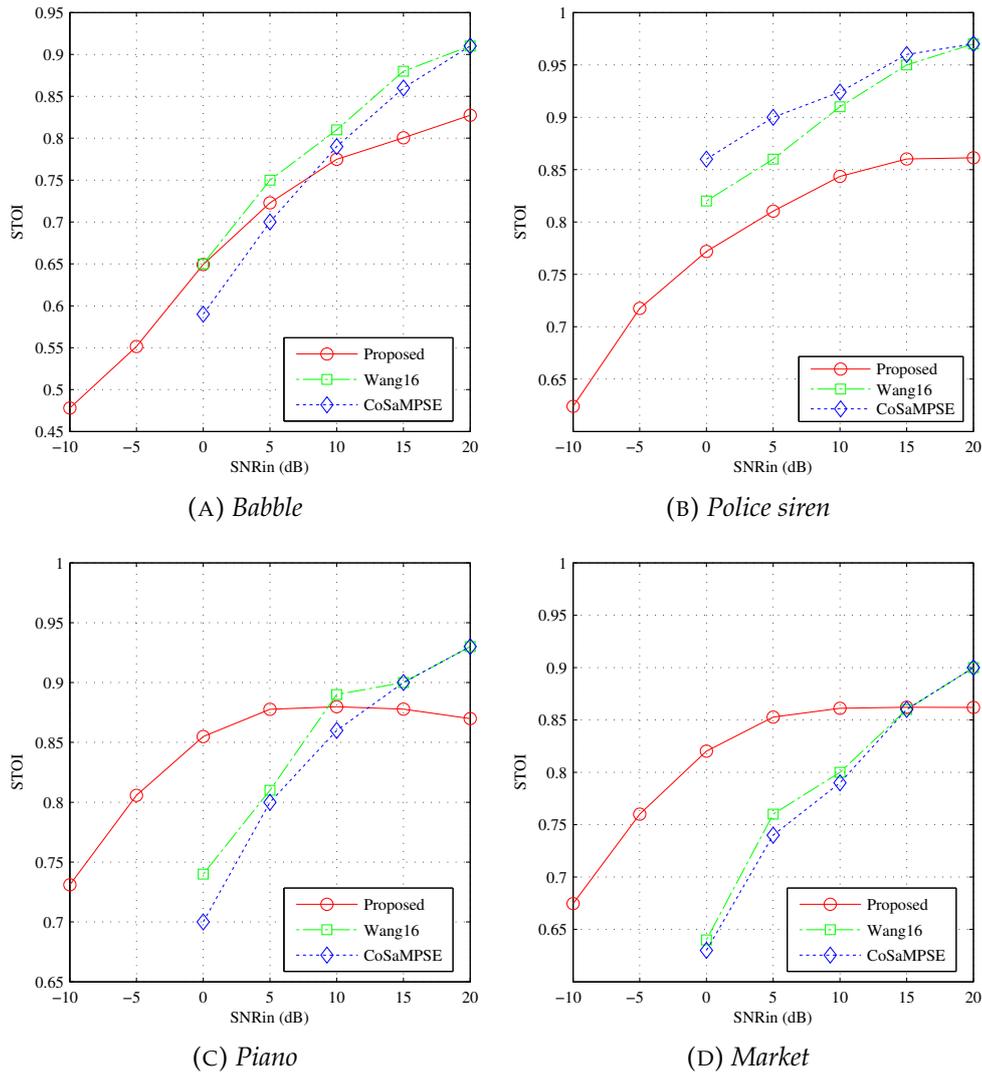


FIGURE 5.5: STOI performances of the proposed CS-based speech enhancement method, Wang16 and CoSaMPSE in various noise conditions: (a) *babble*, (b) *police siren*, (c) *piano*, and (d) *market*. The STOI values of Wang16 and CoSaMPSE are not available for the input SNRs -10 dB and -5 dB.

obtained STOI results using our proposed method significantly outperform those of Wang16 and CoSaMPSE for input SNR levels ≤ 10 dB. For *market* noise (Fig. 5.5 (d)), our STOI scores are higher than those of Wang16 and CoSaMPSE for input SNR levels ≤ 15 dB. In this work, we do not use techniques to enhance the speech intelligibility because we are interested mainly in speech quality enhancement. This fact can justify why the STOI scores are sometimes (e.g., for *police siren* noise (Fig. 5.5 (b))) lower than those of Wang16 and CoSaMPSE.

White and *factory* noises are also considered in our experiments. Fig. 5.6 highlights the comparison between our PESQ scores and those achieved using supervised NMF

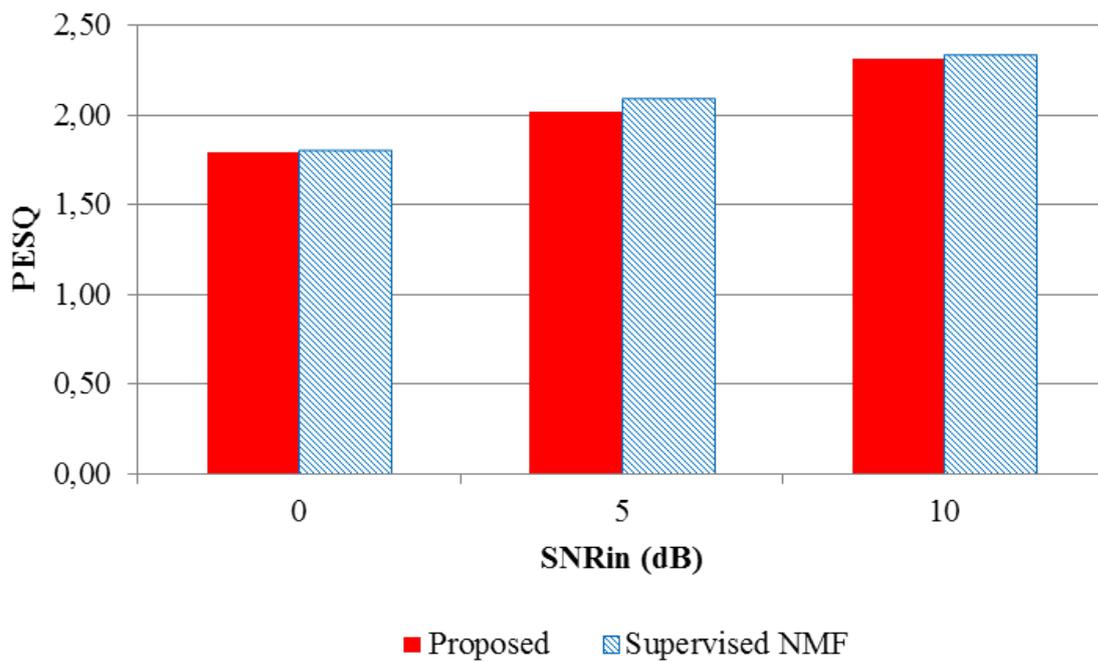


FIGURE 5.6: PESQ performances of the proposed CS-based speech enhancement method and supervised NMF in *white* noise.

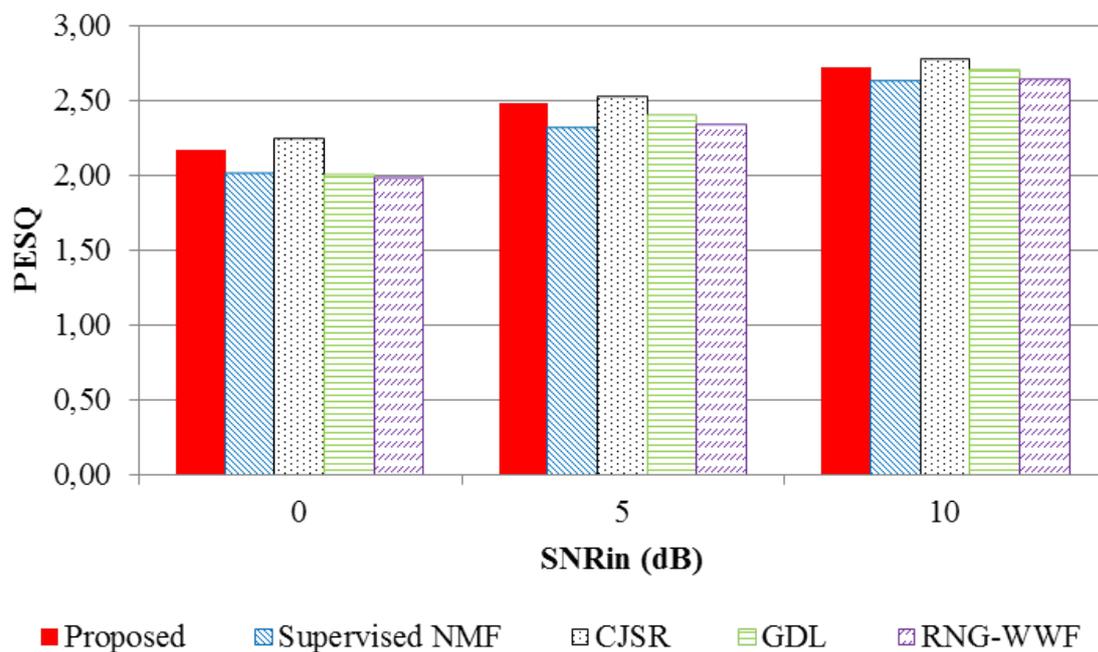


FIGURE 5.7: PESQ performances of the proposed CS-based speech enhancement method, supervised NMF, CJSR, GDL, and RNG-WWF in *factory* noise.

(Mohammadiha, Smaragdīs, and Leijon, 2013) in *white* noise. One can note that the results are comparable at various input SNR levels.

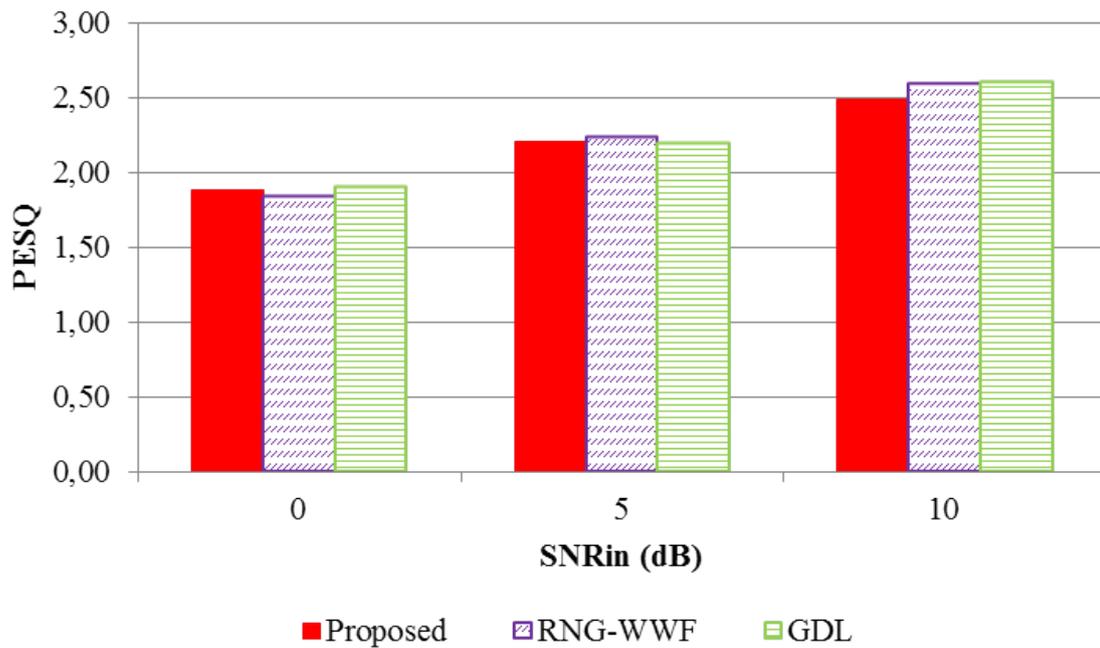


FIGURE 5.8: PESQ performances of the proposed CS-based speech enhancement method, RNG-WWF, and GDL in *babble* noise.

Fig. 5.7 compares our PESQ scores with those obtained using supervised NMF (Mohammadiha, Smaragdis, and Leijon, 2013), CJSR (Luo et al., 2016), GDL (Sigg, Dikk, and Buhmann, 2012), and RNG-WWF (Chung, Plourde, and Champagne, 2017) in *factory* noise. We observe that for various input SNRs, our results are slightly better than those obtained using supervised NMF, GDL, and RNG-WWF, and they are comparable to the scores achieved using CJSR.

The proposed method was also compared with RNG-WWF and GDL for *babble* noise in terms of PESQ scores as shown in Fig. 5.8. The latter demonstrates that our PESQ scores are comparable to those of RNG-WWF and GDL for various input SNR levels.

In Table 5.4, comparison is made, in terms of segmental SNR improvement and PESQ, between our results and those obtained using the CS-based speech enhancement methods proposed by Low (Low13) (Low, Pham, and Venkatesh, 2013), Wang (Wang16) (Wang et al., 2016), and CoSaMPSE. We observe that in terms of segmental SNR, our approach achieves the highest improvements. Additionally, the proposed method gives better PESQ scores than those achieved by Wang16 and CoSaMPSE.

Fig. 5.9 shows representation of clean, noisy, and enhanced speech (for one simulation example). As reported in these graphs, the enhanced signal presents less noise compared to the noisy signal.

TABLE 5.4: Segmental SNR improvements of the proposed CS-based speech enhancement method, Low13, Wang16, and CoSaMPSE in *babble* noise.

Measure	Method	SNRin (dB)				
		-10	-5	0	5	10
SNRseg	Proposed	3.87	3.75	3.64	3.40	3.01
	Low2013	—	1.17	1.25	1.08	1
	Wang16	—	—	3.3	2.8	2
	CoSaMPSE	—	—	0.75	1	0.75
PESQ	Proposed	1.21	1.54	1.89	2.21	2.49
	Low2013	—	1.60	2	2.45	2.75
	Wang16	—	—	1.30	1.70	2
	CoSaMPSE	—	—	1.40	1.70	1.95

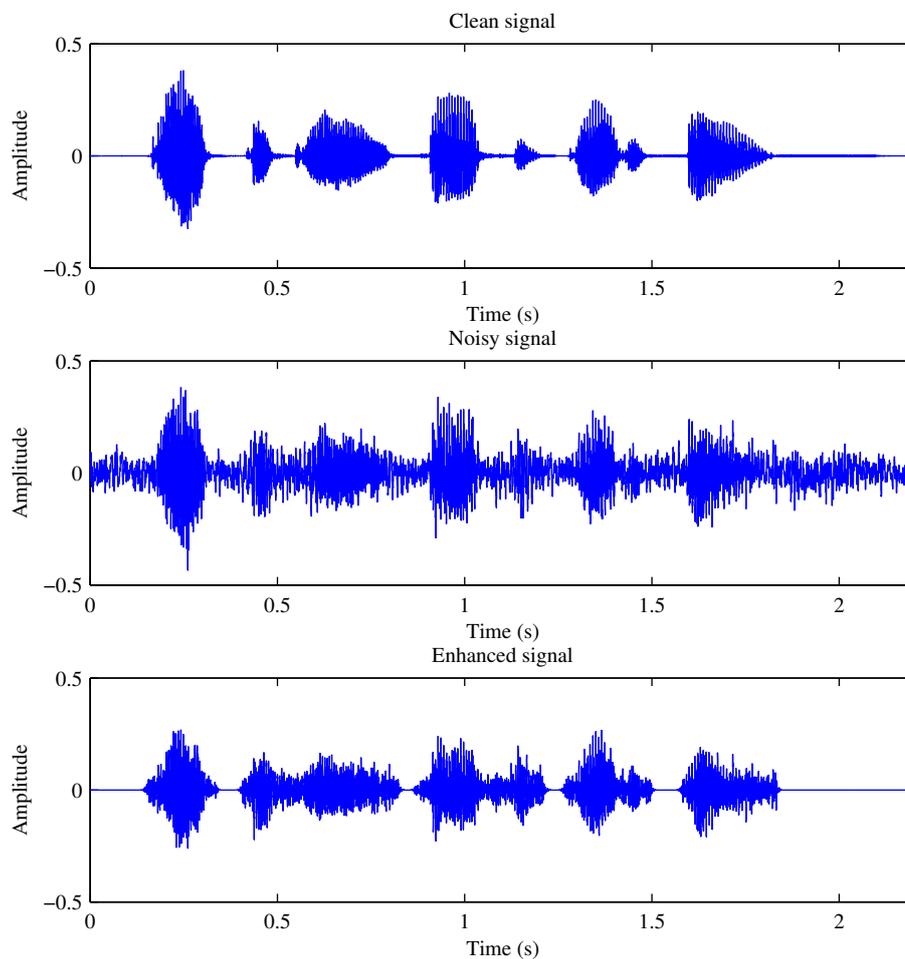


FIGURE 5.9: Clean (top), noisy (middle), and enhanced (bottom) speech waveforms . Note that This signals is the sentence “Wipe the grease off his dirty face” (|'waɪp ðə 'grɪ:s ɔ:f ɪz 'dɜ:ti 'feɪs|) pronounced by a male. The noisy signal is obtained by corrupting the clean speech by babble noise at 0 dB.

From the comparisons above, key findings emerge:

- The proposed CS-based speech enhancement method performs well, giving good results in terms of segmental SNR, PESQ, and STOI scores.
- When comparing our results to those of the previous CS-based speech enhancement methods and other recent techniques (non-based on CS), it must be pointed out that better (and sometimes comparable) performances (in terms of SNRseg and PESQ) are achieved by our approach for different noise types and levels.
- Another promising finding casts light on the fact that for low input SNRs, the proposed method delivers a significant improvement in terms of SNRseg and PESQ scores in various noise conditions (*babble, police siren, piano, and market*) compared to compressive sampling matching pursuit speech enhancement (CoSaMPSE) and Wang16 methods. This is significant because these CS-based speech enhancement methods suffer from the performance decrease in low SNR environments. This can be justified by the fact that
 1. CoSaMPSE performs pure CS on all the signal segments; hence the enhancement performance decreases as it depends on the recovery quality that degrades with the SNR drops;
 2. in Wang16 technique, a mask function is applied on the spectrogram before the sparse recovery to keep only the reliable frequency bins. The purpose is to guarantee that frequency bins dominated by the noise do not affect the CS recovery. Nevertheless, in low SNR, the energy of the noise is large, and the frequency bins that remain after masking are not reliable enough. Hence, noise still has an impact on the sparse recovery performance.

In contrast, our method further reduces the noise effect by performing noise subtraction in the measurement domain, for active speech segments, before the sparse recovery. Moreover, our mask function considers all the frequency bins of silence segments non-reliable. Consequently, the CS reconstruction is applied only on the estimated clean measurement $\hat{\mathbf{Y}}_s$ (see Eq. (5.20)). Hence, the enhancement performance is improved significantly when the SNR is low.

It is worth noting that it is important to choose an efficient VAD technique because errors in the speech/silence detection may influence on the enhanced speech quality.

5.7 Conclusion

In this chapter a novel approach for CS-based speech enhancement is proposed. The new scheme performs noise subtraction in the measurement domain in addition to sparse recovery. Details of the proposed approach are given after providing brief descriptions of three classical speech enhancement methods and the basic CS-based speech enhancement technique.

Results have shown that the proposed CS-based speech enhancement approach has good performances better than those of other recent methods (based/not based on CS), especially for low SNRs. For instance, a segmental SNR improvement of 3.75 dB and a PESQ score of 1.54 are obtained in -5 dB babble noise. As such, the proposed method is a good alternative for speech enhancement.

Chapter 6

Conclusions and Perspectives

The recent advances in applications lead to a huge amount of circulating data in the network, which pose a serious logistical and computational challenges for mobile communication systems. This requires urgent solutions to more efficiently acquire and compress signals. This study set out to exploit the compressed sensing (CS) paradigm for speech coding and communication. This project was undertaken to design a new speech codec based on CS, and to evaluate its behaviour in mobile communication conditions to end up with a novel mobile communication scheme. Furthermore, we had been interested in background noise that may mask the speech signal, and hence give poor performance in many applications, including speech coding and communication. Thus, the second aim of this study was to propose an alternative speech enhancement method based on CS.

In Chapter 3, our idea was to exploit the sparsity of speech in the wavelet basis and the dimensionality reduction property of CS to design a new speech codec based on simply quantizing the CS measurements. We proposed two speech codecs that perform simultaneously the acquisition and compression. The first is based on split vector quantization (SVQ) of compressive sampling measurements, and the second uses split-multistage vector quantization (S-MSVQ). Moreover, new results for speech coding using scalar quantization (SQ), vector quantization, and multistage vector quantization are provided as the latter three schemes were only used for image compression in literature. We also derived the mathematical formula to compute the bitrate when the proposed codecs are used. Results has revealed that speech coding by performing SQ of CS measurements allows *good* quality and intelligibility scores. In fact, we have shown that SQ using four bits per sample does not affect the performance of CS reconstruction, and that this choice leads to trade-off between the bitrate and the speech quality. However, for lower bitrates, our proposed speech codecs (CS with SVQ/S-MSVQ) perform better. Indeed, using the designed CS-based speech codecs, *good* quality performance is achieved with an acceptable execution

time. The best quality performance is achieved when using the proposed codec based on S-MSVQ of CS measurements. Our results are also better than those of other CS-based speech codecs for lower bitrates or compression ratios.

In Chapter 4, a new end-to-end communication system was designed to increase transmission speed, robustness, and security in order to meet the requirements of mobile systems that know an exponentially increasing data amount over time. The design relies on the use of CS-speech codec instead of the supported speech coding standards in actual mobile communication systems, e.g., ITU-T G.722.2. Firstly, we considered a CS-source coding method based on SQ of CS measurements. This allows reducing the speech coding complexity (by using simple quantization and binary encoding), increasing the transmission speed, saving the communication system resources, and encrypting communications without additional costs. Its performance was investigated in worst case communication conditions (Rayleigh channel). Relevant techniques; namely, convolutional coding, OFDM modulation, and diversity are used to mitigate the channel effects; hence, the new secure and robust end-to-end mobile communication scheme. Results have reported that for a bit rate of 12.8 kbit/s the proposed scheme achieves *fair* speech intelligibility and *good* output speech quality scores in 10 dB Rayleigh environment. Secondly, for dealing with lower bitrates, we proposed an improved version of the new mobile system by incorporating the designed speech codec based on S-MSVQ of CS measurements instead of the speech codec based on SQ of CS measurements. Results have reported that for a bit rate of 8.85 kbit/s and in 10 dB Rayleigh environment, the recovered speech has a *good* quality and a *fair* intelligibility scores. Moreover, in degraded communication environments and lower bitrates, the comparison with recent CS-based speech coding methods has shown the merit of the proposed systems.

In Chapter 5, we dealt with the background noise problem, especially at low SNRs where the performance of CS reconstruction decrease. We proposed a novel speech enhancement approach based on CS. Our method performs noise subtraction in the measurement domain in addition to sparse recovery. Dictionary learning, using K-singular value decomposition algorithm, is performed to create an overcomplete dictionary. The noise in the measurement domain is estimated during pauses. Voice activity detection (VAD) is used to classify speech/silence frames. We note here that a good choice of VAD technique is important for an accurate speech/silence detection; otherwise, the enhanced speech quality may be influenced by speech classification errors. Based on the VAD output, a mask function is created, and applied to the noisy speech spectrogram. Furthermore, from each active-speech observation vector, the estimated noise observation vector is subtracted. The enhanced speech spectra are

obtained by sparse recovery using orthogonal matching pursuit. Inverse short-time Fourier transform is used to obtain the enhanced speech. Objective quality and intelligibility evaluations have indicated that our method outperforms the baseline methods in many noisy environments. The proposed CS-based speech enhancement approach achieves good results in low SNRs, for example, a segmental SNR improvement of 3.75 and a PESQ score of 1.54 are obtained in -5 dB *babble* noise.

Our proposed communication systems are designed to meet the various demands in terms of rapidity and security constraints, as in remote control systems and military mobile communications.

Future direction of this research would focus on designing a communication system based on scalable CS-source coding. This suggests that using an adaptive compression ratio depending on the speech segment sparsity and the environment conditions may enhance the quality of the recovered speech at the receiver. Furthermore, considering transmit diversity combined with receive diversity within the proposed communication system (MIMO) would also improve the results.

Index

- ℓ_1 minimization, 108
- ℓ_1 minimization, 10, 11, 21, 22, 52, 93, 105, 116
- Convolutional coding, 81, 87, 88, 91
- CS, compressed sensing, 9, 11, 18, 21, 30, 43, 52, 79, 81, 88, 111, 116, 117, 123
- Diversity, 81, 84, 86, 88, 105
- K-SVD, K-singular value decomposition, 18, 36, 117, 122
- Mobile systems, 42, 79, 81–83, 88, 96, 107
- MRC, maximum ratio combining, 87, 88, 93, 96, 104, 105
- OFDM, orthogonal frequency division multiplexing, 81, 84, 85, 88, 92, 93
- OMP, orthogonal matching pursuit, 18, 23, 24, 36, 118, 121, 123
- PESQ, perceptual evaluation of speech quality, 53, 54, 96, 101, 102, 105, 122–124, 126
- Rayleigh channel, 83, 84, 86, 88, 93
- S-MSVQ, split-multistage vector quantization, 43, 45, 49, 52, 56, 71, 95, 102, 106
- Sparsity, 12, 43, 52, 80, 90, 96, 97, 111, 112, 116, 118, 122
- Speech coding, 35, 42, 43, 45, 48, 49, 56, 81, 88, 106, 111
- Speech enhancement, 35, 111–113, 116, 124
- SQ, scalar quantization, 43, 45, 52, 54, 88, 90, 91, 95
- SVQ, split vector quantization, 43, 45, 48, 52, 56, 68
- VAD, voice activity detection, 35, 112, 118, 120, 123, 124, 133

Bibliography

- Abebe, A. T. and C. G. Kang (2019). "Joint Channel Estimation and MUD for Scalable Grant-free Random Access". In: *IEEE Communications Letters*, pp. 1–1. DOI: [10.1109/LCOMM.2019.2945577](https://doi.org/10.1109/LCOMM.2019.2945577).
- Abrol, V., P. Sharma, and A. K. Sao (2015). "Voiced/nonvoiced detection in compressively sensed speech signals". In: *Speech Communication* 72, pp. 194–207. DOI: [10.1016/j.specom.2015.06.001](https://doi.org/10.1016/j.specom.2015.06.001).
- Ahmed, S. et al. (2015). "Nonparametric Denoising Methods Based on Contourlet Transform with Sharp Frequency Localization: Application to Low Exposure Time Electron Microscopy Images". In: *Entropy* 17.5, pp. 3461–3478. DOI: [10.3390/e17053461](https://doi.org/10.3390/e17053461).
- Akbarpour-Kasgari, A. and M. Ardebilipour (2019). "Massive MIMO-OFDM Channel Estimation via Distributed Compressed Sensing". In: *IEEE Wireless Communications Letters* 8.2, pp. 376–379. DOI: [10.1109/LWC.2018.2873339](https://doi.org/10.1109/LWC.2018.2873339).
- Akl, A., C. Feng, and S. Valaee (2011). "A novel accelerometer-based gesture recognition system". In: *IEEE Transactions on Signal Processing* 59.12, pp. 6197–6205. DOI: [10.1109/tsp.2011.2165707](https://doi.org/10.1109/tsp.2011.2165707).
- Akl, A. and S. Valaee (2010). "Accelerometer-based gesture recognition via dynamic-time warping, affinity propagation, & compressive sensing". In: *2010 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (Dallas, TX, USA). IEEE, pp. 2270–2273. DOI: [10.1109/icassp.2010.5495895](https://doi.org/10.1109/icassp.2010.5495895).
- Al-Azawi, M. K. and A. M. Gaze (2018). "Combined Speech Compression and Encryption using Chaotic Compressive Sensing with Large Key Size". In: *IET Signal Processing* 12.2, pp. 214–218. DOI: [10.1049/iet-spr.2016.0708](https://doi.org/10.1049/iet-spr.2016.0708).
- Alam, M. and Q. Zhang (2018). "Non-Orthogonal Multiple Access With Sequence Block Compressed Sensing Multiuser Detection for 5G". In: *IEEE Access* 6, pp. 63058–63070. DOI: [10.1109/ACCESS.2018.2877477](https://doi.org/10.1109/ACCESS.2018.2877477).
- Almajai, I. and B. Milner (2011). "Visually Derived Wiener Filters for Speech Enhancement". In: *IEEE Transactions on Audio, Speech, and Language Processing* 19.6, pp. 1642–1651. DOI: [10.1109/tasl.2010.2096212](https://doi.org/10.1109/tasl.2010.2096212).
- Alqadah, H. F. (2016). "A Compressive Multi-Frequency Linear Sampling Method for Underwater Acoustic Imaging". In: *IEEE Transactions on Image Processing* 25.6, pp. 2444–2455. DOI: [10.1109/tip.2016.2548243](https://doi.org/10.1109/tip.2016.2548243).
- Amin, M. (2015). *Compressive Sensing for Urban Radar*. en. 1st. Florida, USA: CRC Press, p. 508. ISBN: 978-1-4665-9785-3.
- Andraš, I. et al. (2018). "A time domain reconstruction method of randomly sampled frequency sparse signal". In: *Measurement* 127, pp. 68–77. DOI: <https://doi.org/10.1016/j.measurement.2018.05.065>.
- Andrés, A. M. et al. (2014). "Face recognition on partially occluded images using compressed sensing". In: *Pattern Recognition Letters* 36, pp. 235–242. DOI: [10.1016/j.patrec.2013.08.001](https://doi.org/10.1016/j.patrec.2013.08.001).

- Ariyavisitakul, S., N. R. Sollenberger, and L. J. Greenstein (1997). "Tap-selectable decision-feedback equalization". In: *IEEE Transactions on Communications* 45.12, pp. 1497–1500. DOI: [10.1109/26.650219](https://doi.org/10.1109/26.650219).
- Atia, G. K. and V. Saligrama (2012). "Boolean Compressed Sensing and Noisy Group Testing". In: *IEEE Transactions on Information Theory* 58.3, pp. 1880–1901. DOI: [10.1109/tit.2011.2178156](https://doi.org/10.1109/tit.2011.2178156).
- Başaran, M., S. Erköçük, and H. A. Çırpan (2016). "Bayesian compressive sensing for primary user detection". In: *IET Signal Processing* 10.5, pp. 514–523. DOI: [10.1049/iet-spr.2015.0529](https://doi.org/10.1049/iet-spr.2015.0529).
- Baig, Y., E. M.-K. Lai, and J. P. Lewis (2010). "Quantization effects on Compressed Sensing Video". In: *2010 17th International Conference on Telecommunications (Doha, Qatar)*. IEEE, pp. 935–940. DOI: [10.1109/ictel.2010.5478657](https://doi.org/10.1109/ictel.2010.5478657).
- Bajwa, W. U. et al. (2010). "Compressed Channel Sensing: A New Approach to Estimating Sparse Multipath Channels". In: *Proceedings of the IEEE* 98.6, pp. 1058–1076. DOI: [10.1109/JPROC.2010.2042415](https://doi.org/10.1109/JPROC.2010.2042415).
- Bajwa, W. et al. (2006). "Compressive wireless sensing". In: *Proceedings of the 5th international conference on Information processing in sensor networks (IPSN'06)* (Nashville, Tennessee, USA). ACM Press, pp. 134–142. DOI: [10.1145/1127777.1127801](https://doi.org/10.1145/1127777.1127801).
- Bao, G. et al. (2013). "A Compressed Sensing Approach to Blind Separation of Speech Mixture Based on a Two-Layer Sparsity Model". In: *IEEE Transactions on Audio, Speech, and Language Processing* 21.5, pp. 899–906. DOI: [10.1109/tasl.2012.2234110](https://doi.org/10.1109/tasl.2012.2234110).
- Baraniuk, R. et al. (2008). "A Simple Proof of the Restricted Isometry Property for Random Matrices". In: *Constructive Approximation* 28.3, pp. 253–263. DOI: [10.1007/s00365-007-9003-x](https://doi.org/10.1007/s00365-007-9003-x).
- Baraniuk, R. G. and M. B. Wakin (2007). "Random projections of smooth manifolds". In: *Foundations of computational mathematics* 9.1, pp. 51–77. DOI: [10.1007/s10208-007-9011-z](https://doi.org/10.1007/s10208-007-9011-z).
- Bavarva, A., P. V. Jani, and K. Ghetiya (2018). "Performance Improvement of Wireless Multimedia Sensor Networks Using MIMO and Compressive Sensing". In: *Journal of Communications and Information Networks* 3.1, pp. 84–90. DOI: [10.1007/s41650-018-0011-8](https://doi.org/10.1007/s41650-018-0011-8).
- Bazerque, J. A. and G. B. Giannakis (2010). "Distributed Spectrum Sensing for Cognitive Radio Networks by Exploiting Sparsity". In: *IEEE Transactions on Signal Processing* 58.3, pp. 1847–1862. DOI: [10.1109/TSP.2009.2038417](https://doi.org/10.1109/TSP.2009.2038417).
- Becker, S., J. Bobin, and E. J. Candès (2010). *NESTA: A Fast and Accurate First-Order Method for Sparse Recovery*. <https://statweb.stanford.edu/~candes/nesta/>. v.1.1.
- (2011). "NESTA: A Fast and Accurate First-Order Method for Sparse Recovery". In: *SIAM Journal on Imaging Sciences* 4.1, pp. 1–39. DOI: [10.1137/090756855](https://doi.org/10.1137/090756855).
- Becker, S., E. J. Candès, and M. Grant (2013). *TFOCS: Templates for First-Order Conic Solvers*. <http://cvxr.com/tfocs/>.
- Becker, S. R., E. J. Candès, and M. C. Grant (2011). "Templates for convex cone problems with applications to sparse signal recovery". In: *Mathematical Programming Computation* 3.3, pp. 165–218. DOI: [10.1007/s12532-011-0029-5](https://doi.org/10.1007/s12532-011-0029-5).
- Bencivenni, C. et al. (2015). "Design of Maximally Sparse Antenna Arrays in the Presence of Mutual Coupling". In: *IEEE Antennas and Wireless Propagation Letters* 14, pp. 159–162. DOI: [10.1109/LAWP.2014.2357450](https://doi.org/10.1109/LAWP.2014.2357450).
- (2016). "Synthesis of Maximally Sparse Arrays Using Compressive Sensing and Full-Wave Analysis for Global Earth Coverage Applications". In: *IEEE Transactions on Antennas and Propagation* 64.11, pp. 4872–4877. DOI: [10.1109/TAP.2016.2594840](https://doi.org/10.1109/TAP.2016.2594840).

- Benesty, J., S. Makino, and J. Chen (2005). "Introduction". In: *Speech Enhancement. Signals and Communication Technology*. Signals and Communication Technology. Berlin, Heidelberg: Springer-Verlag, pp. 1–8. ISBN: 978-3-540-27489-6. DOI: [10.1007/3-540-27489-8_1](https://doi.org/10.1007/3-540-27489-8_1).
- Berg, E. van den and M. P. Friedlander (2008). "Probing the Pareto frontier for basis pursuit solutions". In: *SIAM Journal on Scientific Computing* 31.2, pp. 890–912. DOI: [10.1137/080714488](https://doi.org/10.1137/080714488).
- (2018). *SPGL1: A solver for large-scale sparse reconstruction*. <http://www.cs.ubc.ca/labs/sc1/spgl1.v2.0-beta>.
- Berger, C. R. et al. (2010a). "Application of compressive sensing to sparse channel estimation". In: *IEEE Communications Magazine* 48.11, pp. 164–174. DOI: [10.1109/MCOM.2010.5621984](https://doi.org/10.1109/MCOM.2010.5621984).
- Berger, C. R. et al. (2010b). "Sparse Channel Estimation for Multicarrier Underwater Acoustic Communication: From Subspace Methods to Compressed Sensing". In: *IEEE Transactions on Signal Processing* 58.3, pp. 1708–1721. DOI: [10.1109/TSP.2009.2038424](https://doi.org/10.1109/TSP.2009.2038424).
- Berouti, M., R. Schwartz, and J. Makhoul (1979). "Enhancement of speech corrupted by acoustic noise". In: *ICASSP'79. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* (Washington, DC, USA). Vol. 4. IEEE, pp. 208–211. DOI: [10.1109/icassp.1979.1170788](https://doi.org/10.1109/icassp.1979.1170788).
- Bessette, B. et al. (2002). "The adaptive multirate wideband speech codec (AMR-WB)". In: *IEEE Transactions on Speech and Audio Processing* 10.8, pp. 620–636. DOI: [10.1109/tsa.2002.804299](https://doi.org/10.1109/tsa.2002.804299).
- Bhateja, A. K. et al. (2016). "Iris recognition based on sparse representation and k-nearest subspace with genetic algorithm". In: *Pattern Recognition Letters* 73, pp. 13–18. DOI: [10.1016/j.patrec.2015.12.009](https://doi.org/10.1016/j.patrec.2015.12.009).
- Bi, H. et al. (2018). "Complex-Image-Based Sparse SAR Imaging and its Equivalence". In: *IEEE Transactions on Geoscience and Remote Sensing* 56.9, pp. 5006–5014. DOI: [10.1109/tgrs.2018.2803802](https://doi.org/10.1109/tgrs.2018.2803802).
- Bianchi, T., V. Bioglio, and E. Magli (2014). "On the security of random linear measurements". In: *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (Florence, Italy). IEEE, pp. 3992–3996. DOI: [10.1109/icassp.2014.6854351](https://doi.org/10.1109/icassp.2014.6854351).
- (2016). "Analysis of one-time random projections for privacy preserving compressed sensing". In: *IEEE Transactions on Information Forensics and Security* 11.2, pp. 313–327. DOI: [10.1109/tifs.2015.2493982](https://doi.org/10.1109/tifs.2015.2493982).
- Blacknell, D. (2016). "A comparison of compressive sensing and fourier reconstructions for radar target recognition". In: *2016 4th International Workshop on Compressed Sensing Theory and its Applications to Radar, Sonar and Remote Sensing (CoSeRa)* (Aachen, Germany). IEEE, pp. 217–221. DOI: [10.1109/cosera.2016.7745732](https://doi.org/10.1109/cosera.2016.7745732).
- Blumensath, T. and M. Davies (2008). "Gradient Pursuits". In: *IEEE Transactions on Signal Processing* 56.6, pp. 2370–2382. DOI: [10.1109/tsp.2007.916124](https://doi.org/10.1109/tsp.2007.916124).
- (2010). "Normalized Iterative Hard Thresholding: Guaranteed Stability and Performance". In: *IEEE Journal of Selected Topics in Signal Processing* 4.2, pp. 298–309. DOI: [10.1109/jstsp.2010.2042411](https://doi.org/10.1109/jstsp.2010.2042411).
- Blumensath, T. and M. E. Davies (2009). "Iterative hard thresholding for compressed sensing". In: *Applied and Computational Harmonic Analysis* 27.3, pp. 265–274. DOI: [10.1016/j.acha.2009.04.002](https://doi.org/10.1016/j.acha.2009.04.002).
- Boashash, B., ed. (2015). *Time-Frequency Signal Analysis and Processing*. 2nd ed. London, UK: Elsevier & Academic Press.
- Boll, S. (1979). "Suppression of acoustic noise in speech using spectral subtraction". In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 27.2, pp. 113–120. DOI: [10.1109/tassp.1979.1163209](https://doi.org/10.1109/tassp.1979.1163209).

- Bortolotti, D. et al. (2018). "Energy-Aware Bio-Signal Compressed Sensing Reconstruction on the WBSN-Gateway". In: *IEEE Transactions on Emerging Topics in Computing* 6.3, pp. 370–381. DOI: [10.1109/tetc.2016.2564361](https://doi.org/10.1109/tetc.2016.2564361).
- Brungart, D. S. et al. (2006). "Isolating the energetic component of speech-on-speech masking with ideal time-frequency segregation". In: *The Journal of the Acoustical Society of America* 120.6, pp. 4007–4018. DOI: [10.1121/1.2363929](https://doi.org/10.1121/1.2363929).
- Cambareri, V. et al. (2015). "Low-Complexity Multiclass Encryption by Compressed Sensing". In: *IEEE Transactions on Signal Processing* 63.9, pp. 2183–2195. DOI: [10.1109/tsp.2015.2407315](https://doi.org/10.1109/tsp.2015.2407315).
- Candès, E. and J. Romberg (2007). "Sparsity and incoherence in compressive sampling". In: *Inverse Problems* 23.3, pp. 969–985. DOI: [10.1088/0266-5611/23/3/008](https://doi.org/10.1088/0266-5611/23/3/008).
- Candès, E. J. and J. Romberg (2015). *ℓ_1 -MAGIC: Recovery of Sparse Signals via Convex Programming*. <https://statweb.stanford.edu/~candes/l1magic/>.
- Candès, E. J. and M. B. Wakin (2008). "An Introduction To Compressive Sampling". In: *IEEE Signal Processing Magazine* 25.2, pp. 21–30. DOI: [10.1109/MSP.2007.914731](https://doi.org/10.1109/MSP.2007.914731).
- Candès, E., J. Romberg, and T. Tao (2006). "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information". In: *IEEE Transactions on Information Theory* 52.2, pp. 489–509. DOI: [10.1109/tit.2005.862083](https://doi.org/10.1109/tit.2005.862083).
- Candès, E. and T. Tao (2005). "Decoding by Linear Programming". In: *IEEE Transactions on Information Theory* 51.12, pp. 4203–4215. DOI: [10.1109/tit.2005.858979](https://doi.org/10.1109/tit.2005.858979).
- Candès, E. J. and T. Tao (2006). "Near-Optimal Signal Recovery From Random Projections: Universal Encoding Strategies?" In: *IEEE Transactions on Information Theory* 52.12, pp. 5406–5425. DOI: [10.1109/tit.2006.885507](https://doi.org/10.1109/tit.2006.885507).
- Cao, L., T. Hu, and G. Tang (2019). "Shallow-Sea Deghosting via a Compressed Sensing Pseudo-Vertical Velocity Method". In: *IEEE Geoscience and Remote Sensing Letters* 16.3, pp. 357–361. DOI: [10.1109/lgrs.2018.2873666](https://doi.org/10.1109/lgrs.2018.2873666).
- Chan, W. L. et al. (2008). "A single-pixel terahertz imaging system based on compressed sensing". In: *Applied Physics Letters* 93.12, pp. 121105–1–3. DOI: [10.1063/1.2989126](https://doi.org/10.1063/1.2989126).
- Chartrand, R. and V. Staneva (2008). "Restricted isometry properties and nonconvex compressive sensing". In: *Inverse Problems* 24.3, pp. 035020–1–14. DOI: [10.1088/0266-5611/24/3/035020](https://doi.org/10.1088/0266-5611/24/3/035020).
- Chen, C. and J. Wu (2015). "Amplitude-Aided 1-Bit Compressive Sensing Over Noisy Wireless Sensor Networks". In: *IEEE Wireless Communications Letters* 4.5, pp. 473–476. DOI: [10.1109/LWC.2015.2441702](https://doi.org/10.1109/LWC.2015.2441702).
- Chen, J. and X. Huo (2006). "Theoretical Results on Sparse Representations of Multiple-Measurement Vectors". In: *IEEE Transactions on Signal Processing* 54.12, pp. 4634–4643. DOI: [10.1109/TSP.2006.881263](https://doi.org/10.1109/TSP.2006.881263).
- Chen, L., A. Liu, and X. Yuan (2018). "Structured Turbo Compressed Sensing for Massive MIMO Channel Estimation Using a Markov Prior". In: *IEEE Transactions on Vehicular Technology* 67.5, pp. 4635–4639. DOI: [10.1109/TVT.2017.2787708](https://doi.org/10.1109/TVT.2017.2787708).
- Chen, M. et al. (2010). "Compressive sensing on manifolds using a nonparametric mixture of factor analyzers: Algorithm and performance bounds". In: *IEEE Transactions on Signal Processing* 58.12, pp. 6140–6155. DOI: [10.1109/tsp.2010.2070796](https://doi.org/10.1109/tsp.2010.2070796).
- Chen, P. et al. (2017). "Estimation of Extended Targets Based on Compressed Sensing in Cognitive Radar System". In: *IEEE Transactions on Vehicular Technology* 66.2, pp. 941–951. DOI: [10.1109/TVT.2016.2565518](https://doi.org/10.1109/TVT.2016.2565518).
- Chen, S. S., D. L. Donoho, and M. A. Saunders (1998). "Atomic Decomposition by Basis Pursuit". In: *SIAM Journal on Scientific Computing* 20.1, pp. 33–61. DOI: [10.1137/s1064827596304010](https://doi.org/10.1137/s1064827596304010).

- Chen, T. et al. (2019). "Low-Complexity Compressed-Sensing-Based Watermark Cryptosystem and Circuits Implementation for Wireless Sensor Networks". In: *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, pp. 1–13. DOI: [10.1109/TVLSI.2019.2933722](https://doi.org/10.1109/TVLSI.2019.2933722).
- Chen, Y. et al. (2019). "Multi-Resolution Parallel Magnetic Resonance Image Reconstruction in Mobile Computing-Based IoT". In: *IEEE Access* 7, pp. 15623–15633. DOI: [10.1109/access.2019.2894694](https://doi.org/10.1109/access.2019.2894694).
- Chen, Z. et al. (2018). "Efficient and Robust Image Coding and Transmission Based on Scrambled Block Compressive Sensing". In: *IEEE Transactions on Multimedia* 20.7, pp. 1610–1621. DOI: [10.1109/TMM.2017.2774004](https://doi.org/10.1109/TMM.2017.2774004).
- Cheraitia, S.-E. and M. Bouzid (2014). "Robust coding of wideband speech immittance spectral frequencies". In: *Speech Communication* 65, pp. 94–108. DOI: [10.1016/j.specom.2014.07.001](https://doi.org/10.1016/j.specom.2014.07.001).
- Christensen, M. G., J. Ostergaard, and S. H. Jensen (2009). "On compressed sensing and its application to speech and audio signals". In: *2009 Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers (ASILOMAR)* (Pacific Grove, CA, USA). IEEE, pp. 356–360. DOI: [10.1109/acssc.2009.5469828](https://doi.org/10.1109/acssc.2009.5469828).
- Chung, H., E. Plourde, and B. Champagne (2017). "Regularized non-negative matrix factorization with Gaussian mixtures and masking model for speech enhancement". In: *Speech Communication* 87, pp. 18–30. DOI: [10.1016/j.specom.2016.11.003](https://doi.org/10.1016/j.specom.2016.11.003).
- Ciaramella, A. and G. Giunta (2016). "Packet loss recovery in audio multimedia streaming by using compressive sensing". In: *IET Communications* 10.4, pp. 387–392. DOI: [10.1049/iet-com.2014.0995](https://doi.org/10.1049/iet-com.2014.0995).
- Claerbout, J. F. and F. Muir (1973). "Robust Modeling With Erratic Data". In: *Geophysics* 38.5, pp. 826–844. DOI: <http://dx.doi.org/10.1190/1.1440378>.
- Cole, C., M. Karam, and H. Aglan (2008). "Spectral Subtraction of Noise in Speech Processing Applications". In: *2008 40th Southeastern Symposium on System Theory (SSST)* (New Orleans, LA, USA). IEEE, pp. 50–53. DOI: [10.1109/ssst.2008.4480188](https://doi.org/10.1109/ssst.2008.4480188).
- Cormode, G. (2011). "Sketch techniques for approximate query processing". In: *Foundations and Trends in Databases*. NOW publishers.
- Cotter, S. F. and B. D. Rao (2002). "Sparse channel estimation via matching pursuit with application to equalization". In: *IEEE Transactions on Communications* 50.3, pp. 374–377. DOI: [10.1109/26.990897](https://doi.org/10.1109/26.990897).
- Cotter, S. F. et al. (2005). "Sparse solutions to linear inverse problems with multiple measurement vectors". In: *IEEE Transactions on Signal Processing* 53.7, pp. 2477–2488. DOI: [10.1109/TSP.2005.849172](https://doi.org/10.1109/TSP.2005.849172).
- Craven, D. et al. (2014). "Effects of Non-Uniform Quantization on ECG acquired using Compressed Sensing". In: *Proceedings of the 4th International Conference on Wireless Mobile Communication and Healthcare - "Transforming healthcare through innovations in mobile and wireless technologies"* (Athens, Greece). ICST, pp. 1–4. DOI: [10.4108/icst.mobihealth.2014.257357](https://doi.org/10.4108/icst.mobihealth.2014.257357).
- Dai, W. and O. Milenkovic (2009). "Subspace Pursuit for Compressive Sensing Signal Reconstruction". In: *IEEE Transactions on Information Theory* 55.5, pp. 2230–2249. DOI: [10.1109/tit.2009.2016006](https://doi.org/10.1109/tit.2009.2016006).
- (2011). "Information Theoretical and Algorithmic Approaches to Quantized Compressive Sensing". In: *IEEE Transactions on Communications* 59.7, pp. 1857–1866. DOI: [10.1109/tcomm.2011.051711.100204](https://doi.org/10.1109/tcomm.2011.051711.100204).
- Daniels, M. L. and B. D. Rao (2012). "Compressed sensing based scalable speech coders". In: *2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)* (Pacific Grove, CA, USA). IEEE, pp. 92–96. DOI: [10.1109/acssc.2012.6488965](https://doi.org/10.1109/acssc.2012.6488965).
- Daubechies, I., M. Defrise, and C. D. Mol (2004). "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint". In: *Communications on Pure and Applied Mathematics* 57.11, pp. 1413–1457. DOI: [10.1002/cpa.20042](https://doi.org/10.1002/cpa.20042).

- Defraene, B. et al. (2013). "Declipping of Audio Signals Using Perceptual Compressed Sensing". In: *IEEE Transactions on Audio, Speech, and Language Processing* 21.12, pp. 2627–2637. DOI: [10.1109/tasl.2013.2281570](https://doi.org/10.1109/tasl.2013.2281570).
- Donoho, D. L. (2006). "Compressed sensing". In: *IEEE Transactions on Information Theory* 52.4, pp. 1289–1306. DOI: [10.1109/TIT.2006.871582](https://doi.org/10.1109/TIT.2006.871582).
- Donoho, D. L. and B. F. Logan (1992). "Signal Recovery and the Large Sieve". In: *SIAM Journal on Applied Mathematics* 52.2, pp. 577–591. DOI: [10.1137/0152031](https://doi.org/10.1137/0152031).
- Donoho, D. L. and I. M. Johnstone (1994). "Ideal Spatial Adaptation by Wavelet Shrinkage". In: *Biometrika* 81.3, p. 425. DOI: [10.2307/2337118](https://doi.org/10.2307/2337118).
- Donoho, D. L., A. Maleki, and A. Montanari (2009). "Message-passing algorithms for compressed sensing". In: *Proceedings of the National Academy of Sciences* 106.45, pp. 18914–18919. DOI: [10.1073/pnas.0909892106](https://doi.org/10.1073/pnas.0909892106).
- Donoho, D. L. and Y. Tsaig (2008). "Fast Solution of ℓ_1 -Norm Minimization Problems When the Solution May Be Sparse". In: *IEEE Transactions on Information Theory* 54.11, pp. 4789–4812. DOI: [10.1109/tit.2008.929958](https://doi.org/10.1109/tit.2008.929958).
- Donoho, D. (1995). "De-noising by soft-thresholding". In: *IEEE Transactions on Information Theory* 41.3, pp. 613–627. DOI: [10.1109/18.382009](https://doi.org/10.1109/18.382009).
- Donoho, D. and X. Huo (2001). "Uncertainty principles and ideal atomic decomposition". In: *IEEE Transactions on Information Theory* 47.7, pp. 2845–2862. DOI: [10.1109/18.959265](https://doi.org/10.1109/18.959265).
- Donoho, D. et al. (2012). "Sparse Solution of Underdetermined Systems of Linear Equations by Stagewise Orthogonal Matching Pursuit". In: *IEEE Transactions on Information Theory* 58.2, pp. 1094–1121. DOI: [10.1109/TIT.2011.2173241](https://doi.org/10.1109/TIT.2011.2173241).
- Dorfman, R. (1943). "The Detection of Defective Members of Large Populations". In: *The Annals of Mathematical Statistics* 14.4, pp. 436–440. DOI: [10.1214/aoms/1177731363](https://doi.org/10.1214/aoms/1177731363).
- Du, D.-Z. and F. K. Hwang (1993). *Combinatorial Group Testing and Its Applications*. 1st ed. NJ, USA: World Scientific.
- Du, Y. et al. (2018). "Block-Sparsity-Based Multiuser Detection for Uplink Grant-Free NOMA". In: *IEEE Transactions on Wireless Communications* 17.12, pp. 7894–7909. DOI: [10.1109/TWC.2018.2872594](https://doi.org/10.1109/TWC.2018.2872594).
- Duarte, M. F. et al. (2008). "Single-pixel imaging via compressive sampling". In: *IEEE Signal Processing Magazine* 25.2, pp. 83–91. DOI: [10.1109/msp.2007.914730](https://doi.org/10.1109/msp.2007.914730).
- Edgar, M. P. et al. (2016). "Real-time 3D video utilizing a compressed sensing time-of-flight single-pixel camera". In: *Optical Trapping and Optical Micromanipulation XIII* (San Diego, CA, USA). Vol. 9922. SPIE, pp. 171–178. DOI: [10.1117/12.2239113](https://doi.org/10.1117/12.2239113).
- Efron, B. et al. (2004). "Least angle regression". In: *The Annals of Statistics* 32.2, pp. 407–499. DOI: [10.1214/009053604000000067](https://doi.org/10.1214/009053604000000067).
- Elgammal, A., D. Harwood, and L. Davis (2000). "Non-parametric Model for Background Subtraction". In: *Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer, pp. 751–767. DOI: [10.1007/3-540-45053-x_48](https://doi.org/10.1007/3-540-45053-x_48).
- Eltabie, O. M., M. F. Abdelkader, and A. M. Ghuniem (2019). "Incorporating Primary Occupancy Patterns in Compressive Spectrum Sensing". In: *IEEE Access* 7, pp. 29096–29106. DOI: [10.1109/ACCESS.2019.2899953](https://doi.org/10.1109/ACCESS.2019.2899953).
- Eltayeb, M. E., T. Y. Al-Naffouri, and R. W. Heath (2018). "Compressive Sensing for Millimeter Wave Antenna Array Diagnosis". In: *IEEE Transactions on Communications* 66.6, pp. 2708–2721. DOI: [10.1109/TCOMM.2018.2790403](https://doi.org/10.1109/TCOMM.2018.2790403).
- Engan, K., S. O. Aase, and J. H. Husøy (2000). "Multi-frame compression: theory and design". In: *Signal Processing* 80.10, pp. 2121–2140. DOI: [10.1016/s0165-1684\(00\)00072-4](https://doi.org/10.1016/s0165-1684(00)00072-4).

- Fan, X., X. Li, and J. Zhang (2018). "Compressed sensing based loss tomography using weighted ℓ_1 minimization". In: *Computer Communications* 127, pp. 122–130. DOI: <https://doi.org/10.1016/j.comcom.2018.06.004>.
- Farrag, M. (2019). "Secure differential compressive spectrum sensing with 1-bit quantisation". In: *IET Communications* 13.6, pp. 637–641. DOI: [10.1049/iet-com.2018.5279](https://doi.org/10.1049/iet-com.2018.5279).
- Femmam, S., N. M'Sirdi, and A. Ouahabi (2001). "Perception and characterization of materials using signal processing techniques". In: *IEEE Transactions on Instrumentation and Measurement* 50.5, pp. 1203–1211. DOI: [10.1109/19.963184](https://doi.org/10.1109/19.963184).
- Ferroukhi, M. et al. (2019). "Medical Video Coding Based on 2nd-Generation Wavelets: Performance Evaluation". In: *Electronics* 8.1, p. 88. DOI: [10.3390/electronics8010088](https://doi.org/10.3390/electronics8010088).
- Figueiredo, M. A. T., R. D. Nowak, and S. J. Wright (2007). "Gradient Projection for Sparse Reconstruction: Application to Compressed Sensing and Other Inverse Problems". In: *IEEE Journal of Selected Topics in Signal Processing* 1.4, pp. 586–597. DOI: [10.1109/jstsp.2007.910281](https://doi.org/10.1109/jstsp.2007.910281).
- (2009). *GPSR: Gradient Projection for Sparse Reconstruction*. <http://www.lx.it.pt/~mtf/GPSR/>. v.6.0.
- Foucart, S. and H. Rauhut (2013). *A Mathematical Introduction to Compressive Sensing*. NY, USA: Springer.
- Freesound (2019). <http://freesound.org>. Accessed: 2019.02.17.
- Friedman, J., T. Hastie, and R. Tibshirani (2010). "Regularization Paths for Generalized Linear Models via Coordinate Descent". In: *Journal of Statistical Software* 33.1. DOI: [10.18637/jss.v033.i01](https://doi.org/10.18637/jss.v033.i01).
- Frigo, G. (2014). "Compressive Sensing Applications in Measurement: Theoretical issues, algorithm characterization and implementation". eng. PhD Thesis. PhD School in Information Engineering.
- Fu, H. and Y. Chi (2018). "Quantized Spectral Compressed Sensing: Cramer–Rao Bounds and Recovery Algorithms". In: *IEEE Transactions on Signal Processing* 66.12, pp. 3268–3279. DOI: [10.1109/tsp.2018.2827326](https://doi.org/10.1109/tsp.2018.2827326).
- Fuchs, J.-J. (2004). "On Sparse Representations in Arbitrary Redundant Bases". In: *IEEE Transactions on Information Theory* 50.6, pp. 1341–1344. DOI: [10.1109/tit.2004.828141](https://doi.org/10.1109/tit.2004.828141).
- Gan, H., S. Xiao, and Y. Zhao (2018). "A Novel Secure Data Transmission Scheme Using Chaotic Compressed Sensing". In: *IEEE Access* 6, pp. 4587–4598. DOI: [10.1109/ACCESS.2017.2780323](https://doi.org/10.1109/ACCESS.2017.2780323).
- Gao, Z. et al. (2018). "Compressive Sensing Techniques for Next-Generation Wireless Communications". In: *IEEE Wireless Communications* 25.3, pp. 144–153. DOI: [10.1109/mwc.2017.1700147](https://doi.org/10.1109/mwc.2017.1700147).
- Garcia, H., C. V. Correa, and H. Arguello (2018). "Multi-Resolution Compressive Spectral Imaging Reconstruction From Single Pixel Measurements". In: *IEEE Transactions on Image Processing* 27.12, pp. 6174–6184. DOI: [10.1109/tip.2018.2867273](https://doi.org/10.1109/tip.2018.2867273).
- Garofolo, J. S. et al. (1993). *DARPA TIMIT: Acoustic-phonetic Continuous Speech corpus*. Tech. rep. Gaithersburg: NIST.
- Gemmeke, J. F. et al. (2010). "Compressive Sensing for Missing Data Imputation in Noise Robust Speech Recognition". In: *IEEE Journal of Selected Topics in Signal Processing* 4.2, pp. 272–287. DOI: [10.1109/jstsp.2009.2039171](https://doi.org/10.1109/jstsp.2009.2039171).
- George, S. N., N. Augustine, and D. P. Pattathil (2014). "Audio security through compressive sampling and cellular automata". In: *Multimedia Tools and Applications* 74.23, pp. 10393–10417. DOI: [10.1007/s11042-014-2172-2](https://doi.org/10.1007/s11042-014-2172-2).
- Ghosh, R, B Chatterjee, and S Chakravor (2016). "A low-complexity method based on compressed sensing for long term field measurement of insulator leakage current". In: *IEEE Transactions on Dielectrics and Electrical Insulation* 23.1, pp. 596–604. DOI: [10.1109/TDEI.2015.005458](https://doi.org/10.1109/TDEI.2015.005458).
- Giacobello, D. et al. (2010). "Retrieving Sparse Patterns Using a Compressed Sensing Framework: Applications to Speech Coding Based on Sparse Linear Prediction". In: *IEEE Signal Processing Letters* 17.1, pp. 103–106. DOI: [10.1109/lsp.2009.2034560](https://doi.org/10.1109/lsp.2009.2034560).

- Giacobello, D. et al. (2012). "Sparse Linear Prediction and Its Applications to Speech Processing". In: *IEEE Transactions on Audio, Speech, and Language Processing* 20.5, pp. 1644–1657. DOI: [10.1109/tasl.2012.2186807](https://doi.org/10.1109/tasl.2012.2186807).
- Gilbert, A. C., S. Muthukrishnan, and M. J. Strauss (2003). "Approximation of functions over redundant dictionaries using coherence". In: *Proceedings of the fourteenth annual ACM-SIAM symposium on Discrete algorithms (SODA'03)* (Philadelphia, PA, USA). SIAM, pp. 243–252.
- Gishkori, S. and G. Leus (2013). "Compressive Sampling Based Energy Detection of Ultra-Wideband Pulse Position Modulation". In: *IEEE Transactions on Signal Processing* 61.15, pp. 3866–3879. DOI: [10.1109/TSP.2013.2260747](https://doi.org/10.1109/TSP.2013.2260747).
- Gluskin, E. D. (1984). "Norms OF Random Matrices And Widths OF Finite-Dimensional Sets". In: *Mathematics of the USSR-Sbornik* 48.1, pp. 173–182. DOI: [10.1070/sm1984v048n01abeh002667](https://doi.org/10.1070/sm1984v048n01abeh002667).
- Gómez-Cuba, F. and A. J. Goldsmith (2019). "Compressed Sensing Channel Estimation for OFDM with non-Gaussian Multipath Gains". In: *IEEE Transactions on Wireless Communications* 19.1, pp. 47–61. DOI: [10.1109/twc.2019.2941192](https://doi.org/10.1109/twc.2019.2941192).
- Gong, T., Z. Yang, and M. Zheng (2019). "Compressive Subspace Learning Based Wideband Spectrum Sensing for Multiantenna Cognitive Radio". In: *IEEE Transactions on Vehicular Technology* 68.7, pp. 6636–6648. DOI: [10.1109/TVT.2019.2915269](https://doi.org/10.1109/TVT.2019.2915269).
- Grais, E. M. and H. Erdogan (2013). "Regularized nonnegative matrix factorization using Gaussian mixture priors for supervised single channel source separation". In: *Computer Speech & Language* 27.3, pp. 746–762. DOI: [10.1016/j.cs1.2012.09.002](https://doi.org/10.1016/j.cs1.2012.09.002).
- Grant, M. and S. Boyd (2008). "Graph implementations for nonsmooth convex programs". In: *Recent Advances in Learning and Control*. Ed. by V. Blondel, S. Boyd, and H. Kimura. Lecture Notes in Control and Information Sciences. http://stanford.edu/~boyd/graph_dcp.html. Springer-Verlag Limited, pp. 95–110.
- (2018). *CVX: Matlab Software for Disciplined Convex Programming*. <http://cvxr.com/cvx/>. v.2.1.
- Gribonval, R. and M. Nielsen (2003). "Sparse representations in unions of bases". In: *IEEE Transactions on Information Theory* 49.12, pp. 3320–3325. DOI: [10.1109/tit.2003.820031](https://doi.org/10.1109/tit.2003.820031).
- Griffin, A. et al. (2011). "Single-Channel and Multi-Channel Sinusoidal Audio Coding Using Compressed Sensing". In: *IEEE Transactions on Audio, Speech, and Language Processing* 19.5, pp. 1382–1395. DOI: [10.1109/tas1.2010.2090656](https://doi.org/10.1109/tas1.2010.2090656).
- Guo, L. and A. M. Abbosh (2015). "Microwave Stepped Frequency Head Imaging Using Compressive Sensing With Limited Number of Frequency Steps". In: *IEEE Antennas and Wireless Propagation Letters* 14, pp. 1133–1136. DOI: [10.1109/lawp.2015.2396054](https://doi.org/10.1109/lawp.2015.2396054).
- Gurbuz, A. C., J. H. McClellan, and V. Cevher (2008). "A compressive beamforming method". In: *2008 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (Las Vegas, NV, USA). (IEEE), pp. 2617–2620. DOI: [10.1109/ICASSP.2008.4518185](https://doi.org/10.1109/ICASSP.2008.4518185).
- Hale, E. T., W. Yin, and Y. Zhang (2008). "Fixed-Point Continuation for ℓ_1 -Minimization: Methodology and Convergence". In: *SIAM Journal on Optimization* 19.3, pp. 1107–1130. DOI: [10.1137/070698920](https://doi.org/10.1137/070698920).
- Haltmeier, M. et al. (2016). "Compressed sensing and sparsity in photoacoustic tomography". In: *Journal of Optics* 18.11, p. 114004. DOI: [10.1088/2040-8978/18/11/114004](https://doi.org/10.1088/2040-8978/18/11/114004).
- Hamdaoui, B., B. Khalfi, and M. Guizani (2018). "Compressed Wideband Spectrum Sensing: Concept, Challenges, and Enablers". In: *IEEE Communications Magazine* 56.4, pp. 136–141. DOI: [10.1109/MCOM.2018.1700719](https://doi.org/10.1109/MCOM.2018.1700719).
- Han, Z., H. Li, and W. Yin (2013). *Compressive sensing for wireless networks*. Cambridge, UK: Cambridge University Press, p. 293. ISBN: 978-1-107-01883-9.

- Haneche, H., B. Boudraa, and A. Ouahabi (2018a). "Speech Enhancement Using Compressed Sensing-based method". In: *2018 International Conference on Electrical Sciences and Technologies in Maghreb (CISTEM)* (Algiers, Algeria). IEEE, pp. 1–6. DOI: [10.1109/cistem.2018.8613609](https://doi.org/10.1109/cistem.2018.8613609).
- (2018b). "Split Vector Quantization of Compressive Sampling Measurements for Speech Compression". In: *2018 International Conference on Signal, Image, Vision and their Applications (SIVA)* (Guelma, Algeria). IEEE, pp. 1–6. DOI: [10.1109/siva.2018.8661151](https://doi.org/10.1109/siva.2018.8661151).
- (2020). "A new way to enhance speech signal based on compressed sensing". In: *Measurement* 151, p. 107117. DOI: [10.1016/j.measurement.2019.107117](https://doi.org/10.1016/j.measurement.2019.107117).
- Haneche, H., A. Ouahabi, and B. Boudraa (2019a). "Compressed Sensing-Speech Coding Scheme for Mobile Communications". In: *Circuits, Systems, and Signal Processing*. Accepted.
- (2019b). "New mobile communication system design for Rayleigh environments based on compressed sensing-source coding". In: *IET Communications* 13.15, pp. 2375–2385. DOI: [10.1049/iet-com.2018.5348](https://doi.org/10.1049/iet-com.2018.5348).
- Harris, P. et al. (2016). "Monitoring Anthropogenic Ocean Sound from Shipping Using an Acoustic Sensor Network and a Compressive Sensing Approach". In: *Sensors* 16.3, p. 415. DOI: [10.3390/s16030415](https://doi.org/10.3390/s16030415).
- Haupt, J. et al. (2008). "Compressed Sensing for Networked Data". In: *IEEE Signal Processing Magazine* 25.2, pp. 92–101. DOI: [10.1109/MSP.2007.914732](https://doi.org/10.1109/MSP.2007.914732).
- Havary-Nassab, V., S. Hassan, and S. Valaee (2010). "Compressive detection for wide-band spectrum sensing". In: *2010 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (Dallas, TX, USA). IEEE, pp. 3094–3097. DOI: [10.1109/ICASSP.2010.5496101](https://doi.org/10.1109/ICASSP.2010.5496101).
- Hawes, M. et al. (2017). "Bayesian Compressive Sensing Approaches for Direction of Arrival Estimation With Mutual Coupling Effects". In: *IEEE Transactions on Antennas and Propagation* 65.3, pp. 1357–1368. DOI: [10.1109/TAP.2017.2655013](https://doi.org/10.1109/TAP.2017.2655013).
- Hayashi, K., M. Nagahara, and T. Tanaka (2013). "A User's Guide to Compressed Sensing for Communications Systems". In: *IEICE Transactions on Communications* E96.B.3, pp. 685–712. DOI: [10.1587/transcom.E96.B.685](https://doi.org/10.1587/transcom.E96.B.685).
- He, J. et al. (2019). "Compressive Multi-attribute Data Gathering using Hankel Matrix in Wireless Sensor Networks". In: *IEEE Communications Letters* 23.12, pp. 2417–2421. DOI: [10.1109/LCOMM.2019.2941194](https://doi.org/10.1109/LCOMM.2019.2941194).
- He, K. et al. (2018). "A novel compressed sensing-based non-orthogonal multiple access scheme for massive MTC in 5G systems". In: *EURASIP Journal on Wireless Communications and Networking* 2018.81, pp. 1–14. DOI: [10.1186/s13638-018-1079-4](https://doi.org/10.1186/s13638-018-1079-4).
- Hegde, C., M. B. Wakin, and R. G. Baraniuk (2007). "Random Projections for Manifold Learning". In: *Proceedings of the 20th International Conference on Neural Information Processing Systems (NIPS'07)* (Vancouver, British Columbia, Canada). Curran Associates Inc., 641–648.
- Herrmann, F. J. and G. Hennenfent (2008). "Non-parametric seismic data recovery with curvelet frames". In: *Geophysical Journal International* 173.1, pp. 233–248. DOI: [10.1111/j.1365-246x.2007.03698.x](https://doi.org/10.1111/j.1365-246x.2007.03698.x).
- Hou, B. et al. (2017). "Sparse Coding-Inspired High-Resolution ISAR Imaging Using Multistage Compressive Sensing". In: *IEEE Transactions on Aerospace and Electronic Systems* 53.1, pp. 26–40. DOI: [10.1109/taes.2017.2649161](https://doi.org/10.1109/taes.2017.2649161).
- Hoyer, P. O. (2004). "Non-negative matrix factorization with sparseness constraints". In: *Journal of machine learning research* 5.Nov, pp. 1457–1469.
- Hu, H. et al. (2017). "Locating the few: Sparsity-aware waveform design for active radar". In: *IEEE Transactions on Signal Processing* 65.3, pp. 651–662. DOI: [10.1109/tsp.2016.2620966](https://doi.org/10.1109/tsp.2016.2620966).

- Hu, Y. and P. C. Loizou (2007). "Subjective comparison and evaluation of speech enhancement algorithms". In: *Speech Communication* 49.7-8, pp. 588–601. DOI: [10.1016/j.specom.2006.12.006](https://doi.org/10.1016/j.specom.2006.12.006).
- Hu, Y. and P. C. Loizou (2008). "Evaluation of Objective Quality Measures for Speech Enhancement". In: *IEEE Transactions on Audio, Speech, and Language Processing* 16.1, pp. 229–238. DOI: [10.1109/tasl.2007.911054](https://doi.org/10.1109/tasl.2007.911054).
- Huang, C. and L. Wang (2012). "Dynamic Sampling Rate Adjustment for Compressive Spectrum Sensing over Cognitive Radio Network". In: *IEEE Wireless Communications Letters* 1.2, pp. 57–60. DOI: [10.1109/WCL.2012.010912.110136](https://doi.org/10.1109/WCL.2012.010912.110136).
- Huang, H. et al. (2016). "A compressive-sensing based testing vehicle for 3D TSV pre-bond and post-bond testing data". In: *Proceedings of the 2016 on International Symposium on Physical Design (ISPD'16)* (Santa Rosa, CA, USA). ACM Press, pp. 19–25. DOI: [10.1145/2872334.2872351](https://doi.org/10.1145/2872334.2872351).
- Hurley, N. and S. Rickard (2009). "Comparing Measures of Sparsity". In: *IEEE Transactions on Information Theory* 55.10, pp. 4723–4741. DOI: [10.1109/tit.2009.2027527](https://doi.org/10.1109/tit.2009.2027527).
- Ince, T. and G. Ögücü (2016). "Array Failure Diagnosis Using Nonconvex Compressed Sensing". In: *IEEE Antennas and Wireless Propagation Letters* 15, pp. 992–995. DOI: [10.1109/LAWP.2015.2489760](https://doi.org/10.1109/LAWP.2015.2489760).
- Jančovič, P., X. Zou, and M. Köküer (2012). "Speech enhancement based on Sparse Code Shrinkage employing multiple speech models". In: *Speech Communication* 54.1, pp. 108–118. DOI: [10.1016/j.specom.2011.07.005](https://doi.org/10.1016/j.specom.2011.07.005).
- Ji, Y., W.-P. Zhu, and B. Champagne (2019). "Recurrent Neural Network-Based Dictionary Learning for Compressive Speech Sensing". In: *Circuits, Systems, and Signal Processing* 38.8. DOI: [10.1007/s00034-019-01058-5](https://doi.org/10.1007/s00034-019-01058-5).
- Jiang, D. et al. (2018). "Impaired Array Diagnosis and Mitigation With Khatri–Rao Processing". In: *IEEE Antennas and Wireless Propagation Letters* 17.12, pp. 2354–2358. DOI: [10.1109/LAWP.2018.2874802](https://doi.org/10.1109/LAWP.2018.2874802).
- Joneidi, M. et al. (2018). "E-optimal Sensor Selection for Compressive Sensing-based Purposes". In: *IEEE Transactions on Big Data*, pp. 1–1. DOI: [10.1109/TBDATA.2018.2868120](https://doi.org/10.1109/TBDATA.2018.2868120).
- Kamath, S. and P. Loizou (2002). "A multi-band spectral subtraction method for enhancing speech corrupted by colored noise". In: *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)* (Orlando, FL, USA). IEEE, pp. IV–4164–IV–4164. DOI: [10.1109/icassp.2002.5745591](https://doi.org/10.1109/icassp.2002.5745591).
- Kang, M.-S. and K.-T. Kim (2019). "Compressive Sensing Based SAR Imaging and Autofocus Using Improved Tikhonov Regularization". In: *IEEE Sensors Journal* 19.14, pp. 5529–5540. DOI: [10.1109/jsen.2019.2904611](https://doi.org/10.1109/jsen.2019.2904611).
- Kang, M.-S. et al. (2017). "ISAR Imaging of High-Speed Maneuvering Target Using Gapped Stepped-Frequency Waveform and Compressive Sensing". In: *IEEE Transactions on Image Processing* 26.10, pp. 5043–5056. DOI: [10.1109/tip.2017.2728182](https://doi.org/10.1109/tip.2017.2728182).
- Karvanen, J. and A. Cichocki (2003). "Measuring sparseness of noisy signals". In: *4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA)* (Nara, Japan). Citeseer, pp. 125–130.
- Kašin, B. S. (1977). "Diameters of some finite-dimensional sets and classes of smooth functions". In: *Mathematics of the USSR-Izvestiya* 11.2, pp. 317–333. DOI: [10.1070/IM1977v011n02ABEH001719](https://doi.org/10.1070/IM1977v011n02ABEH001719).
- Kates, J. M. and K. H. Arehart (2005). "Coherence and the speech intelligibility index". In: *The Journal of the Acoustical Society of America* 117.4, pp. 2224–2237. DOI: [10.1121/1.1862575](https://doi.org/10.1121/1.1862575).
- Katzberg, F. et al. (2018). "A compressed sensing framework for dynamic sound-field measurements". In: *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 26.11, pp. 1962–1975. DOI: [10.1109/TASLP.2018.2851144](https://doi.org/10.1109/TASLP.2018.2851144).

- Kim, J. M., O. K. Lee, and J. C. Ye (2012). "Compressive MUSIC: Revisiting the Link Between Compressive Sensing and Array Signal Processing". In: *IEEE Transactions on Information Theory* 58.1, pp. 278–301. DOI: [10.1109/TIT.2011.2171529](https://doi.org/10.1109/TIT.2011.2171529).
- Kinsho, H. et al. (2019). "Heterogeneous Delay Tomography for Wide-Area Mobile Networks". In: *IEEE Transactions on Communications* E102.B.8, pp. 1607–1616. DOI: [10.1587/transcom.2018EBP3056](https://doi.org/10.1587/transcom.2018EBP3056).
- Kjems, U. et al. (2009). "Role of mask pattern in intelligibility of ideal binary-masked noisy speech". In: *The Journal of the Acoustical Society of America* 126.3, pp. 1415–1426. DOI: [10.1121/1.3179673](https://doi.org/10.1121/1.3179673).
- Kocic, M., D. Brady, and M. Stojanovic (1995). "Sparse equalization for real-time digital underwater acoustic communications". In: 'Challenges of Our Changing Global Environment'. *Conference Proceedings. OCEANS '95 MTS/IEEE* (San Diego, CA, USA). Vol. 3. IEEE, pp. 1417–1422. DOI: [10.1109/oceans.1995.528671](https://doi.org/10.1109/oceans.1995.528671).
- Kotel'nikov, V. A. (1933). "On the Transmission Capacity of the "Ether" and Wire in Electrocommunications". In: *Material for the First All-Union Conference on Questions of Communication*. Izd. Red. Upr. Svyazi RKKA, Moscow (Russian). English translation by V. E. Katsnelson. DOI: [10.1007/978-1-4612-0143-4_2](https://doi.org/10.1007/978-1-4612-0143-4_2).
- Krahmer, F., R. Saab, and O. Yilmaz (2014). "Sigma-Delta quantization of sub-Gaussian frame expansions and its application to compressed sensing". In: *Information and Inference* 3.1, pp. 40–58. DOI: [10.1093/imaiai/iat007](https://doi.org/10.1093/imaiai/iat007).
- La Rédaction (2018). *Un rapport international le confirme : 50% de la population algérienne utilise Internet*. Ed. by Algérie Part. <https://algeriepart.com/2018/02/26/rapport-international-confirme-50-de-population-algerienne-utilise-internet/>. Accessed: 2019.11.18.
- Laska, J. N. and R. G. Baraniuk (2012). "Regime Change: Bit-Depth Versus Measurement-Rate in Compressive Sensing". In: *IEEE Transactions on Signal Processing* 60.7, pp. 3496–3505. DOI: [10.1109/tsp.2012.2194710](https://doi.org/10.1109/tsp.2012.2194710).
- Lee, H., C. A. Gong, and P. Chen (2019). "A Compressed Sensing Estimation Technique for Doubly Selective Channel in OFDM Systems". In: *IEEE Access* 7, pp. 115192–115199. DOI: [10.1109/ACCESS.2019.2935758](https://doi.org/10.1109/ACCESS.2019.2935758).
- Lei, Z. et al. (2015). "Localization of low-frequency coherent sound sources with compressive beamforming-based passive synthetic aperture". In: *The Journal of the Acoustical Society of America* 137.4, EL255–EL260. DOI: [10.1121/1.4915003](https://doi.org/10.1121/1.4915003).
- Li, C. et al. (2019). "Robust Multichannel EEG Compressed Sensing in the Presence of Mixed Noise". In: *IEEE Sensors Journal* 19.22, pp. 10574–10583. DOI: [10.1109/jsen.2019.2930546](https://doi.org/10.1109/jsen.2019.2930546).
- Li, G. and R. J. Burkholder (2015). "Hybrid matching pursuit for distributed through-wall radar imaging". In: *IEEE Transactions on Antennas and Propagation* 63.4, pp. 1701–1711. DOI: [10.1109/tap.2015.2398115](https://doi.org/10.1109/tap.2015.2398115).
- Li, J. et al. (2016). "Enhanced GPR Signal for Layered Media Time-Delay Estimation in Low-SNR Scenario". In: *IEEE Geoscience and Remote Sensing Letters* 13.3, pp. 299–303. DOI: [10.1109/lgrs.2015.2502662](https://doi.org/10.1109/lgrs.2015.2502662).
- Li, L. et al. (2019). "Flexible and Secure Data Transmission System Based on Semitensor Compressive Sensing in Wireless Body Area Networks". In: *IEEE Internet of Things Journal* 6.2, pp. 3212–3227. DOI: [10.1109/JIOT.2018.2881129](https://doi.org/10.1109/JIOT.2018.2881129).
- Li, L. et al. (2020). "Efficient and Secure Image Communication System Based on Compressed Sensing for IoT Monitoring Applications". In: *IEEE Transactions on Multimedia* 22.1, pp. 82–95. DOI: [10.1109/tmm.2019.2923111](https://doi.org/10.1109/tmm.2019.2923111).

- Li, L. (2014). "Subwavelength Imaging of Sparse Broadband Sources in an Open Disordered Medium From a Single Antenna". In: *IEEE Antennas and Wireless Propagation Letters* 13, pp. 1461–1464. DOI: [10.1109/lawp.2014.2341231](https://doi.org/10.1109/lawp.2014.2341231).
- Li, W., W. Deng, and M. D. Migliore (2018). "A Deterministic Far-Field Sampling Strategy for Array Diagnosis Using Sparse Recovery". In: *IEEE Antennas and Wireless Propagation Letters* 17.7, pp. 1261–1265. DOI: [10.1109/LAWP.2018.2841650](https://doi.org/10.1109/LAWP.2018.2841650).
- Li, X. et al. (2018a). "Millimeter Wave Channel Estimation via Exploiting Joint Sparse and Low-Rank Structures". In: *IEEE Transactions on Wireless Communications* 17.2, pp. 1123–1133. DOI: [10.1109/TWC.2017.2776108](https://doi.org/10.1109/TWC.2017.2776108).
- Li, Y. and Y. Liang (2018). "Compressed Sensing in Multi-Hop Large-Scale Wireless Sensor Networks Based on Routing Topology Tomography". In: *IEEE Access* 6, pp. 27637–27650. DOI: [10.1109/ACCESS.2018.2834550](https://doi.org/10.1109/ACCESS.2018.2834550).
- Li, Z. et al. (2018b). "Dynamic Compressive Wide-Band Spectrum Sensing Based on Channel Energy Reconstruction in Cognitive Internet of Things". In: *IEEE Transactions on Industrial Informatics* 14.6, pp. 2598–2607. DOI: [10.1109/TII.2018.2797096](https://doi.org/10.1109/TII.2018.2797096).
- Liao, C. et al. (2016). "Efficient Spatial Variation Modeling of Nanoscale Integrated Circuits Via Hidden Markov Tree". In: *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 35.6, pp. 971–984. DOI: [10.1109/tcad.2015.2481868](https://doi.org/10.1109/tcad.2015.2481868).
- Lilis, G. N., D. Angelosante, and G. B. Giannakis (2010). "Sound Field Reproduction using the Lasso". In: *IEEE Transactions on Audio, Speech, and Language Processing* 18.8, pp. 1902–1912. DOI: [10.1109/TASL.2010.2040523](https://doi.org/10.1109/TASL.2010.2040523).
- Linde, Y., A. Buzo, and R. Gray (1980). "An algorithm for vector quantizer design". In: *IEEE Transactions on Communications* 28.1, pp. 84–95. DOI: [10.1109/tcom.1980.1094577](https://doi.org/10.1109/tcom.1980.1094577).
- Liu, B. et al. (2018). "Dimension-Reduced Direction-of-Arrival Estimation Based on $\ell_{2,1}$ -Norm Penalty". In: *IEEE Access* 6, pp. 44433–44444. DOI: [10.1109/ACCESS.2018.2862435](https://doi.org/10.1109/ACCESS.2018.2862435).
- Liu, Z., A. Y. Elezzabi, and H. V. Zhao (2011). "Maximum frame rate video acquisition using adaptive compressed sensing". In: *IEEE Transactions on Circuits and Systems for Video Technology* 21.11, pp. 1704–1718. DOI: [10.1109/tcsvt.2011.2133890](https://doi.org/10.1109/tcsvt.2011.2133890).
- Lloyd, S. (1982). "Least squares quantization in PCM". In: *IEEE Transactions on Information Theory* 28.2, pp. 129–137. DOI: [10.1109/tit.1982.1056489](https://doi.org/10.1109/tit.1982.1056489).
- Logan, B. F. (1965). "Properties of high-pass signals". PhD thesis. Columbia University.
- Loizou, P. (2007). *Speech Enhancement: Theory and Practice*. 2nd ed. FL, USA: CRC, Boca Raton.
- Low, S. Y., D. S. Pham, and S. Venkatesh (2013). "Compressive speech enhancement". In: *Speech Communication* 55.6, pp. 757–768. DOI: [10.1016/j.specom.2013.03.003](https://doi.org/10.1016/j.specom.2013.03.003).
- Lu, Y. and P. C. Loizou (2011). "Estimators of the Magnitude-Squared Spectrum and Methods for Incorporating SNR Uncertainty". In: *IEEE Transactions on Audio, Speech, and Language Processing* 19.5, pp. 1123–1137. DOI: [10.1109/tas1.2010.2082531](https://doi.org/10.1109/tas1.2010.2082531).
- Luo, Y. et al. (2016). "Supervised Monaural Speech Enhancement Using Complementary Joint Sparse Representations". In: *IEEE Signal Processing Letters* 23.2, pp. 237–241. DOI: [10.1109/lsp.2015.2509480](https://doi.org/10.1109/lsp.2015.2509480).
- Mairal, J. (2017). *SPAMS: SParse Modeling Software*. <http://spams-devel.gforge.inria.fr>. v.2.6.
- Mairal, J. et al. (2009). "Online dictionary learning for sparse coding". In: *Proceedings of the 26th Annual International Conference on Machine Learning - ICML'09 (Monreal, Canada)*. ACM Press, pp. 689–696. DOI: [10.1145/1553374.1553463](https://doi.org/10.1145/1553374.1553463).
- (2010). "Online learning for matrix factorization and sparse coding". In: *Journal of Machine Learning Research* 11, Jan, pp. 19–60.

- Malioutov, D., M. Cetin, and A. Willsky (2005). "Homotopy continuation for sparse signal representation". In: *ICASSP'05. IEEE International Conference on Acoustics, Speech, and Signal Processing* (Philadelphia, PA). Vol. 5. IEEE, pp. v/733–v/736. DOI: [10.1109/icassp.2005.1416408](https://doi.org/10.1109/icassp.2005.1416408).
- Mallat, S. and Z. Zhang (1993). "Matching pursuits with time-frequency dictionaries". In: *IEEE Transactions on Signal Processing* 41.12, pp. 3397–3415. DOI: [10.1109/78.258082](https://doi.org/10.1109/78.258082).
- Mamaghanian, H. et al. (2011). "Compressed Sensing for Real-Time Energy-Efficient ECG Compression on Wireless Body Sensor Nodes". In: *IEEE Transactions on Biomedical Engineering* 58.9, pp. 2456–2466. DOI: [10.1109/tbme.2011.2156795](https://doi.org/10.1109/tbme.2011.2156795).
- Mammone, N. et al. (2019). "Brain Network Analysis of Compressive Sensed High-Density EEG Signals in AD and MCI Subjects". In: *IEEE Transactions on Industrial Informatics* 15.1, pp. 527–536. DOI: [10.1109/tii.2018.2868431](https://doi.org/10.1109/tii.2018.2868431).
- Mangia, M. et al. (2019). "Chained Compressed Sensing: A Blockchain-Inspired Approach for Low-Cost Security in IoT Sensing". In: *IEEE Internet of Things Journal* 6.4, pp. 6465–6475. DOI: [10.1109/JIOT.2019.2910402](https://doi.org/10.1109/JIOT.2019.2910402).
- Mardani, M. et al. (2019). "Deep Generative Adversarial Neural Networks for Compressive Sensing MRI". In: *IEEE Transactions on Medical Imaging* 38.1, pp. 167–179. DOI: [10.1109/tmi.2018.2858752](https://doi.org/10.1109/tmi.2018.2858752).
- Mathew, R. S. and J. S. Paul (2018). "Sparsity Promoting Adaptive Regularization for Compressed Sensing Parallel MRI". In: *IEEE Transactions on Computational Imaging* 4.1, pp. 147–159. DOI: [10.1109/tci.2017.2787911](https://doi.org/10.1109/tci.2017.2787911).
- Matsuda, T., M. Nagahara, and K. Hayashi (2011). "Link Quality Classifier with Compressed Sensing Based on $\ell_1 - \ell_2$ Optimization". In: *IEEE Communications Letters* 15.10, pp. 1117–1119. DOI: [10.1109/LCOMM.2011.082911.111611](https://doi.org/10.1109/LCOMM.2011.082911.111611).
- Migliore, M. D. (2011). "A Compressed Sensing Approach for Array Diagnosis From a Small Set of Near-Field Measurements". In: *IEEE Transactions on Antennas and Propagation* 59.6, pp. 2127–2133. DOI: [10.1109/TAP.2011.2144556](https://doi.org/10.1109/TAP.2011.2144556).
- Mohammadiha, N., P. Smaragdis, and A. Leijon (2013). "Supervised and Unsupervised Speech Enhancement Using Nonnegative Matrix Factorization". In: *IEEE Transactions on Audio, Speech, and Language Processing* 21.10, pp. 2140–2151. DOI: [10.1109/tas1.2013.2270369](https://doi.org/10.1109/tas1.2013.2270369).
- Morabito, A. F. (2017). "Synthesis of Maximum-Efficiency Beam Arrays via Convex Programming and Compressive Sensing". In: *IEEE Antennas and Wireless Propagation Letters* 16, pp. 2404–2407. DOI: [10.1109/LAWP.2017.2721218](https://doi.org/10.1109/LAWP.2017.2721218).
- Mourad, R., Z. Dawy, and F. Morcos (2013). "Designing Pooling Systems for Noisy High-Throughput Protein-Protein Interaction Experiments Using Boolean Compressed Sensing". In: *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 10.6, pp. 1478–1490. DOI: [10.1109/tcbb.2013.129](https://doi.org/10.1109/tcbb.2013.129).
- Mun, S. and J. E. Fowler (2012). "DPCM for quantized block-based compressed sensing of images". In: *2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO)* (Bucharest, Romania). IEEE, pp. 1424–1428.
- Nagesh, P. and B. Li (2009a). "A compressive sensing approach for expression-invariant face recognition". In: *2009 IEEE Conference on Computer Vision and Pattern Recognition* (Miami, FL, USA). IEEE, pp. 1518–1525. DOI: [10.1109/cvpr.2009.5206657](https://doi.org/10.1109/cvpr.2009.5206657).
- (2009b). "Compressive imaging of color images". In: *2009 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (Taipei, Taiwan). IEEE, pp. 1261–1264. DOI: [10.1109/icassp.2009.4959820](https://doi.org/10.1109/icassp.2009.4959820).
- Nakanishi, K. et al. (2014). "Synchronization-Free Delay Tomography Based on Compressed Sensing". In: *IEEE Communications Letters* 18.8, pp. 1343–1346. DOI: [10.1109/LCOMM.2014.2334303](https://doi.org/10.1109/LCOMM.2014.2334303).

- Nakanishi, K. et al. (2018). "Route referencing and ordering for synchronization-free delay tomography in wireless networks". In: *EURASIP Journal on Wireless Communications and Networking* 2018.1, p. 211. DOI: [10.1186/s13638-018-1227-x](https://doi.org/10.1186/s13638-018-1227-x).
- Naseem, I., R. Togneri, and M. Bennamoun (2010). "Sparse representation for speaker identification". In: *2010 20th International Conference on Pattern Recognition (Istanbul, Turkey)*. IEEE, pp. 4460–4463. DOI: [10.1109/icpr.2010.1083](https://doi.org/10.1109/icpr.2010.1083).
- Natarajan, B. K. (1995). "Sparse Approximate Solutions to Linear Systems". In: *SIAM Journal on Computing* 24.2, pp. 227–234. DOI: [10.1137/s0097539792240406](https://doi.org/10.1137/s0097539792240406).
- Needell, D. and J. Tropp (2009). "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples". In: *Applied and Computational Harmonic Analysis* 26.3, pp. 301–321. DOI: [10.1016/j.acha.2008.07.002](https://doi.org/10.1016/j.acha.2008.07.002).
- Needell, D. and R. Vershynin (2009). "Uniform Uncertainty Principle and Signal Recovery via Regularized Orthogonal Matching Pursuit". In: *Foundations of Computational Mathematics* 9.3, pp. 317–334. DOI: [10.1007/s10208-008-9031-3](https://doi.org/10.1007/s10208-008-9031-3).
- Nguyen, L. H. and T. D. Tran (2016). "Efficient and Robust RFI Extraction Via Sparse Recovery". In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 9.6, pp. 2104–2117. DOI: [10.1109/JSTARS.2016.2528884](https://doi.org/10.1109/JSTARS.2016.2528884).
- Nyquist, H. (1928). "Certain Topics in Telegraph Transmission Theory". In: *Transactions of the American Institute of Electrical Engineers* 47.2, pp. 617–644. DOI: [10.1109/t-aiee.1928.5055024](https://doi.org/10.1109/t-aiee.1928.5055024).
- Oliveri, G. and A. Massa (2011). "Bayesian Compressive Sampling for Pattern Synthesis With Maximally Sparse Non-Uniform Linear Arrays". In: *IEEE Transactions on Antennas and Propagation* 59.2, pp. 467–481. DOI: [10.1109/TAP.2010.2096400](https://doi.org/10.1109/TAP.2010.2096400).
- Orsdemir, A. et al. (2008). "On the security and robustness of encryption via compressed sensing". In: *MILCOM 2008 - 2008 IEEE Military Communications Conference (San Diego, CA, USA)*. IEEE, pp. 1–7. DOI: [10.1109/milcom.2008.4753187](https://doi.org/10.1109/milcom.2008.4753187).
- Osborne, M. R., B. Presnell, and B. A. Turlach (2000). "On the LASSO and its Dual". In: *Journal of Computational and Graphical Statistics* 9.2, pp. 319–337. DOI: [10.1080/10618600.2000.10474883](https://doi.org/10.1080/10618600.2000.10474883).
- Ouahabi, A. (2012). *Signal and Image Multiresolution Analysis*. 1st ed. London-Hoboken, UK-USA: ISTE-John Wiley & Sons.
- Ouahabi, A. (2013). "A review of wavelet denoising in medical imaging". In: *2013 8th International Workshop on Systems, Signal Processing and their Applications (WoSSPA) (Algiers, Algeria)*. IEEE. DOI: [10.1109/wosspa.2013.6602330](https://doi.org/10.1109/wosspa.2013.6602330).
- P.862.2 (2007). *Wideband extension to Recommendation P.862 for the assessment of wideband telephone networks and speech codecs*. Geneva, Switzerland.
- Paliwal, K., K. Wójcicki, and B. Schwerin (2010). "Single-channel speech enhancement using spectral subtraction in the short-time modulation domain". In: *Speech Communication* 52.5, pp. 450–475. DOI: [10.1016/j.specom.2010.02.004](https://doi.org/10.1016/j.specom.2010.02.004).
- Palmeri, R., T. Isernia, and A. F. Morabito (2019). "Diagnosis of Planar Arrays Through Phaseless Measurements and Sparsity Promotion". In: *IEEE Antennas and Wireless Propagation Letters* 18.6, pp. 1273–1277. DOI: [10.1109/LAWP.2019.2914529](https://doi.org/10.1109/LAWP.2019.2914529).
- Paredes, J. L., G. R. Arce, and Z. Wang (2007). "Ultra-Wideband Compressed Sensing: Channel Estimation". In: *IEEE Journal of Selected Topics in Signal Processing* 1.3, pp. 383–395. DOI: [10.1109/JSTSP.2007.906657](https://doi.org/10.1109/JSTSP.2007.906657).
- Parvaresh, F. et al. (2008). "Recovering Sparse Signals Using Sparse Measurement Matrices in Compressed DNA Microarrays". In: *IEEE Journal of Selected Topics in Signal Processing* 2.3, pp. 275–285. DOI: [10.1109/jstsp.2008.924384](https://doi.org/10.1109/jstsp.2008.924384).

- Pati, Y., R. Rezaifar, and P. Krishnaprasad (1993). "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition". In: *Proceedings of 27th Asilomar Conference on Signals, Systems and Computers* (Pacific Grove, CA, USA). IEEE Comput. Soc. Press, pp. 40–44. DOI: [10.1109/acssc.1993.342465](https://doi.org/10.1109/acssc.1993.342465).
- Pelissier, M. and C. Studer (2018). "Non-Uniform Wavelet Sampling for RF Analog-to-Information Conversion". In: *IEEE Transactions on Circuits and Systems I: Regular Papers* 65.2, pp. 471–484. DOI: [10.1109/TCSI.2017.2729779](https://doi.org/10.1109/TCSI.2017.2729779).
- Peng, H. et al. (2017). "Secure and Energy-Efficient Data Transmission System Based on Chaotic Compressive Sensing in Body-to-Body Networks". In: *IEEE Transactions on Biomedical Circuits and Systems* 11.3, pp. 558–573. DOI: [10.1109/TBCAS.2017.2665659](https://doi.org/10.1109/TBCAS.2017.2665659).
- Peng, X. et al. (2014). "Incorporating reference in parallel imaging and compressed sensing". In: *Magnetic Resonance in Medicine* 73.4, pp. 1490–1504. DOI: [10.1002/mrm.25272](https://doi.org/10.1002/mrm.25272).
- Pieraccini, M. and L. Miccinesi (2019). "An Interferometric MIMO Radar for Bridge Monitoring". In: *IEEE Geoscience and Remote Sensing Letters* 16.9, pp. 1383–1387. DOI: [10.1109/lgrs.2019.2900405](https://doi.org/10.1109/lgrs.2019.2900405).
- Potts, D. and M. Tasche (2010). "Parameter estimation for exponential sums by approximate Prony method". In: *Signal Processing* 90.5, pp. 1631–1642. DOI: [10.1016/j.sigpro.2009.11.012](https://doi.org/10.1016/j.sigpro.2009.11.012).
- Prisco, G. and M. D'Urso (2012). "Maximally Sparse Arrays Via Sequential Convex Optimizations". In: *IEEE Antennas and Wireless Propagation Letters* 11, pp. 192–195. DOI: [10.1109/LAWP.2012.2186626](https://doi.org/10.1109/LAWP.2012.2186626).
- Prony, R. (1795). "Essai experimental et analytique". In: *J. de l'Ecole Polytechnique* 2, pp. 24–35.
- Pu, W. et al. (2019). "Fast Compressive Sensing-Based SAR Imaging Integrated With Motion Compensation". In: *IEEE Access* 7, pp. 53284–53295. DOI: [10.1109/access.2019.2911696](https://doi.org/10.1109/access.2019.2911696).
- Pudlewski, S. and T. Melodia (2013a). "A Tutorial on Encoding and Wireless Transmission of Compressively Sampled Videos". In: *IEEE Communications Surveys Tutorials* 15.2, pp. 754–767. DOI: [10.1109/SURV.2012.121912.00154](https://doi.org/10.1109/SURV.2012.121912.00154).
- (2013b). "Compressive Video Streaming: Design and Rate-Energy-Distortion Analysis". In: *IEEE Transactions on Multimedia* 15.8, pp. 2072–2086. DOI: [10.1109/TMM.2013.2280245](https://doi.org/10.1109/TMM.2013.2280245).
- Pudlewski, S., A. Prasanna, and T. Melodia (2012). "Compressed-Sensing-Enabled Video Streaming for Wireless Multimedia Sensor Networks". In: *IEEE Transactions on Mobile Computing* 11.6, pp. 1060–1072. DOI: [10.1109/TMC.2011.175](https://doi.org/10.1109/TMC.2011.175).
- Qi, C. and L. Wu (2011). "Optimized Pilot Placement for Sparse Channel Estimation in OFDM Systems". In: *IEEE Signal Processing Letters* 18.12, pp. 749–752. DOI: [10.1109/LSP.2011.2170834](https://doi.org/10.1109/LSP.2011.2170834).
- Qi, J. et al. (2015). "A Hybrid Security and Compressive Sensing-Based Sensor Data Gathering Scheme". In: *IEEE Access* 3, pp. 718–724. DOI: [10.1109/ACCESS.2015.2439034](https://doi.org/10.1109/ACCESS.2015.2439034).
- Qian, C., X. Fu, and N. D. Sidiropoulos (2019). "Amplitude Retrieval for Channel Estimation of MIMO Systems with One-Bit ADCs". In: *IEEE Signal Processing Letters* 26.11, pp. 1698–1702. DOI: [10.1109/LSP.2019.2945490](https://doi.org/10.1109/LSP.2019.2945490).
- Qin, T. et al. (2015). "Quality Improvement of Thermoacoustic Imaging Based on Compressive Sensing". In: *IEEE Antennas and Wireless Propagation Letters* 14, pp. 1200–1203. DOI: [10.1109/lawp.2015.2397952](https://doi.org/10.1109/lawp.2015.2397952).
- Qing, L., H. Guangyao, and F. Xiaomei (2018). "Physical Layer Security in Multi-Hop AF Relay Network Based on Compressed Sensing". In: *IEEE Communications Letters* 22.9, pp. 1882–1885. DOI: [10.1109/LCOMM.2018.2853101](https://doi.org/10.1109/LCOMM.2018.2853101).
- Qiu, W., J. Zhou, and Q. Fu (2020). "Jointly Using Low-Rank and Sparsity Priors for Sparse Inverse Synthetic Aperture Radar Imaging". In: *IEEE Transactions on Image Processing* 29, pp. 100–115. DOI: [10.1109/tip.2019.2927458](https://doi.org/10.1109/tip.2019.2927458).

- Qu, L. et al. (2015). "Diffraction Tomographic Ground-Penetrating Radar Multibistatic Imaging Algorithm With Compressive Frequency Measurements". In: *IEEE Geoscience and Remote Sensing Letters* 12.10, pp. 2011–2015. DOI: [10.1109/lgrs.2015.2441991](https://doi.org/10.1109/lgrs.2015.2441991).
- Quan, T. M., T. Nguyen-Duc, and W.-K. Jeong (2018). "Compressed Sensing MRI Reconstruction Using a Generative Adversarial Network With a Cyclic Loss". In: *IEEE Transactions on Medical Imaging* 37.6, pp. 1488–1497. DOI: [10.1109/tmi.2018.2820120](https://doi.org/10.1109/tmi.2018.2820120).
- Ramdas, V., D. Mishra, and S. S. Gorthi (2015). "Speech coding and enhancement using quantized compressive sensing measurements". In: *2015 IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES)* (Kozhikode, India). IEEE, pp. 1–5. DOI: [10.1109/spices.2015.7091436](https://doi.org/10.1109/spices.2015.7091436).
- Rao, B. and K. Kreutz-Delgado (1999). "An affine scaling methodology for best basis selection". In: *IEEE Transactions on Signal Processing* 47.1, pp. 187–200. DOI: [10.1109/78.738251](https://doi.org/10.1109/78.738251).
- Rickard, S. and M. Fallon (2004). "The Gini index of speech". In: *Proceedings of the 38th Conference on Information Science and Systems (CISS'04)* (Edinburgh, Scotland, UK), pp. 1–5.
- Rousset, F. et al. (2017). "Adaptive Basis Scan by Wavelet Prediction for Single-Pixel Imaging". In: *IEEE Transactions on Computational Imaging* 3.1, pp. 36–46. DOI: [10.1109/tci.2016.2637079](https://doi.org/10.1109/tci.2016.2637079).
- Rudin, L. I., S. Osher, and E. Fatemi (1992). "Nonlinear total variation based noise removal algorithms". In: *Physica D: Nonlinear Phenomena* 60.1-4, pp. 259–268. DOI: [10.1016/0167-2789\(92\)90242-f](https://doi.org/10.1016/0167-2789(92)90242-f).
- Ruyet, D. L. and M. Pischella (2015). *Digital Communications 1: Source and Channel Coding*. 2nd ed. London-Hoboken, UK-USA: ISTE-John Wiley & Sons.
- Salucci, M. et al. (2018). "Planar Array Diagnosis by Means of an Advanced Bayesian Compressive Processing". In: *IEEE Transactions on Antennas and Propagation* 66.11, pp. 5892–5906. DOI: [10.1109/TAP.2018.2866534](https://doi.org/10.1109/TAP.2018.2866534).
- Santosa, F. and W. W. Symes (1986). "Linear Inversion of Band-Limited Reflection Seismograms". In: *SIAM Journal on Scientific and Statistical Computing* 7.4, pp. 1307–1330. DOI: [10.1137/0907087](https://doi.org/10.1137/0907087).
- Sekkate, S., M. Khalil, and A. Adib (2019). "Speaker Identification for OFDM-Based Aeronautical Communication System". In: *Circuits, Systems, and Signal Processing*. ISSN: 1531-5878. DOI: [10.1007/s00034-019-01026-z](https://doi.org/10.1007/s00034-019-01026-z).
- Seyedtabaee, S. and H. M. Goodarzi (2010). "Improved Noise Minimum Statistics Estimation Algorithm for Using in a Speech-Passing Noise-Rejecting Headset". In: *EURASIP Journal on Advances in Signal Processing* 2010.1, pp. 1–11. DOI: [10.1155/2010/395048](https://doi.org/10.1155/2010/395048).
- Shannon, C. E. (1948). "A Mathematical Theory of Communication". In: *Bell System Technical Journal* 27.4, pp. 623–656. DOI: [10.1002/j.1538-7305.1948.tb00917.x](https://doi.org/10.1002/j.1538-7305.1948.tb00917.x).
- Shannon, C. (1949). "Communication in the Presence of Noise". In: *Proceedings of the IRE* 37.1, pp. 10–21. DOI: [10.1109/jrproc.1949.232969](https://doi.org/10.1109/jrproc.1949.232969).
- Sharma, S., A. Gupta, and V. Bhatia (2016). "A New Sparse Signal-Matched Measurement Matrix for Compressive Sensing in UWB Communication". In: *IEEE Access* 4, pp. 5327–5342. DOI: [10.1109/ACCESS.2016.2601779](https://doi.org/10.1109/ACCESS.2016.2601779).
- (2019). "Compressed Sensing Based UWB Receiver Using Signal-Matched Sparse Measurement Matrix". In: *IEEE Transactions on Vehicular Technology* 68.1, pp. 993–998. DOI: [10.1109/TVT.2018.2881509](https://doi.org/10.1109/TVT.2018.2881509).
- Sharma, S. K. et al. (2016). "Application of Compressive Sensing in Cognitive Radio Communications: A Survey". In: *IEEE Communications Surveys Tutorials* 18.3, pp. 1838–1860. DOI: [10.1109/COMST.2016.2524443](https://doi.org/10.1109/COMST.2016.2524443).
- Shi, F. (2019). "Two Dimensional Direction-of-Arrival Estimation Using Compressive Measurements". In: *IEEE Access* 7, pp. 20863–20868. DOI: [10.1109/ACCESS.2019.2892085](https://doi.org/10.1109/ACCESS.2019.2892085).

- Shi, G. et al. (2008). "UWB Echo Signal Detection With Ultra-Low Rate Sampling Based on Compressed Sensing". In: *IEEE Transactions on Circuits and Systems II: Express Briefs* 55.4, pp. 379–383. DOI: [10.1109/TCSII.2008.918988](https://doi.org/10.1109/TCSII.2008.918988).
- Shim, B. and B. Song (2012). "Multiuser Detection via Compressive Sensing". In: *IEEE Communications Letters* 16.7, pp. 972–974. DOI: [10.1109/LCOMM.2012.050112.111980](https://doi.org/10.1109/LCOMM.2012.050112.111980).
- Shirazinia, A., S. Chatterjee, and M. Skoglund (2012). "Performance bounds for vector quantized compressive sensing". In: *2012 International Symposium on Information Theory and its Applications* (Honolulu, HI, USA). IEEE, pp. 289–293.
- Shirazinia, A., S. Chatterjee, and M. Skoglund (2013). "Analysis-by-Synthesis Quantization for Compressed Sensing Measurements". In: *IEEE Transactions on Signal Processing* 61.22, pp. 5789–5800. DOI: [10.1109/tsp.2013.2280445](https://doi.org/10.1109/tsp.2013.2280445).
- (2014). "Joint Source-Channel Vector Quantization for Compressed Sensing". In: *IEEE Transactions on Signal Processing* 6.14, pp. 3667–3681. DOI: [10.1109/tsp.2014.2329649](https://doi.org/10.1109/tsp.2014.2329649).
- Sigg, C. D., T. Dikk, and J. M. Buhmann (2012). "Speech Enhancement Using Generative Dictionary Learning". In: *IEEE Transactions on Audio, Speech, and Language Processing* 20.6, pp. 1698–1712. DOI: [10.1109/tasl.2012.2187194](https://doi.org/10.1109/tasl.2012.2187194).
- Sivapalan, S. et al. (2011). "Compressive sensing for gait recognition". In: *2011 International Conference on Digital Image Computing: Techniques and Applications* (Noosa, QLD, Australia). IEEE, pp. 567–571. DOI: [10.1109/dicta.2011.101](https://doi.org/10.1109/dicta.2011.101).
- Sivaram, G. S. et al. (2010). "Sparse coding for speech recognition". In: *2010 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (Dallas, TX, USA). IEEE, pp. 4346–4349. DOI: [10.1109/icassp.2010.5495649](https://doi.org/10.1109/icassp.2010.5495649).
- Sklar, B. (2001). *Digital Communications: Fundamentals and Applications*. 2nd ed. New Jersey, USA: Prentice Hall PTR.
- Sreenivas, T. V. and W. B. Kleijn (2009). "Compressive sensing for sparsely excited speech signals". In: *2009 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (Taipei, Taiwan). IEEE, pp. 4125–4128. DOI: [10.1109/icassp.2009.4960536](https://doi.org/10.1109/icassp.2009.4960536).
- Stankovic, L. and M. Brajovic (2018). "Analysis of the Reconstruction of Sparse Signals in the DCT Domain Applied to Audio Signals". In: *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)* 26.7, pp. 1216–1231. DOI: [10.1109/TASLP.2018.2819819](https://doi.org/10.1109/TASLP.2018.2819819).
- Stüber, G. L. (2017). *Principles of Mobile Communication*. 4th ed. New York, USA: Springer International Publishing. DOI: [10.1007/978-3-319-55615-4](https://doi.org/10.1007/978-3-319-55615-4).
- Sun, H. et al. (2018). "Kirchhoff Beam Migration Based on Compressive Sensing". In: *IEEE Access* 6, pp. 26520–26529. DOI: [10.1109/access.2018.2828160](https://doi.org/10.1109/access.2018.2828160).
- Sun, J. Z. and V. K. Goyal (2009). "Optimal quantization of random measurements in compressed sensing". In: *2009 IEEE International Symposium on Information Theory* (Seoul, South Korea). IEEE, pp. 6–10. DOI: [10.1109/isit.2009.5205695](https://doi.org/10.1109/isit.2009.5205695).
- Sun, L. et al. (2019). "A Deep Information Sharing Network for Multi-Contrast Compressed Sensing MRI Reconstruction". In: *IEEE Transactions on Image Processing* 28.12, pp. 6141–6153. DOI: [10.1109/tip.2019.2925288](https://doi.org/10.1109/tip.2019.2925288).
- Sun, Y. et al. (2016). "MT-BCS-Based Two-Dimensional Diffraction Tomographic GPR Imaging Algorithm With Multiview–Multistatic Configuration". In: *IEEE Geoscience and Remote Sensing Letters* 13.6, pp. 831–835. DOI: [10.1109/lgrs.2016.2549538](https://doi.org/10.1109/lgrs.2016.2549538).
- Taal, C. H. et al. (2011). "An Algorithm for Intelligibility Prediction of Time–Frequency Weighted Noisy Speech". In: *IEEE Transactions on Audio, Speech, and Language Processing* 19.7, pp. 2125–2136. DOI: [10.1109/tasl.2011.2114881](https://doi.org/10.1109/tasl.2011.2114881).

- Tang, V. H., A. Bouzerdoum, and S. L. Phung (2018). "Multipolarization Through-Wall Radar Imaging Using Low-Rank and Jointly-Sparse Representations". In: *IEEE Transactions on Image Processing* 27.4, pp. 1763–1776. DOI: [10.1109/tip.2017.2786462](https://doi.org/10.1109/tip.2017.2786462).
- Tashan, T. and M. Al-Azawi (2018). "Multilevel magnetic resonance imaging compression using compressive sensing". In: *IET Image Processing* 12.12, pp. 2186–2191. DOI: [10.1049/iet-ipr.2018.5611](https://doi.org/10.1049/iet-ipr.2018.5611).
- Taubock, G. et al. (2010). "Compressive Estimation of Doubly Selective Channels in Multicarrier Systems: Leakage Effects and Sparsity-Enhancing Processing". In: *IEEE Journal of Selected Topics in Signal Processing* 4.2, pp. 255–271. DOI: [10.1109/JSTSP.2010.2042410](https://doi.org/10.1109/JSTSP.2010.2042410).
- Tibshirani, R. (1996). "Regression shrinkage and selection via the lasso". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 58.1, pp. 267–288. DOI: [10.1111/j.2517-6161.1996.tb02080.x](https://doi.org/10.1111/j.2517-6161.1996.tb02080.x).
- Torruella, P. et al. (2016). "3D visualization of the iron oxidation state in FeO/Fe₃O₄ core-shell nanocubes from electron energy loss tomography". In: *Nano letters* 16.8, pp. 5068–5073. DOI: [10.1021/acs.nanolett.6b01922](https://doi.org/10.1021/acs.nanolett.6b01922).
- Tropp, J. A. and A. C. Gilbert (2007). "Signal Recovery From Random Measurements Via Orthogonal Matching Pursuit". In: *IEEE Transactions on Information Theory* 53.12, pp. 4655–4666. DOI: [10.1109/TIT.2007.909108](https://doi.org/10.1109/TIT.2007.909108).
- Tropp, J. (2004). "Greed is Good: Algorithmic Results for Sparse Approximation". In: *IEEE Transactions on Information Theory* 50.10, pp. 2231–2242. DOI: [10.1109/tit.2004.834793](https://doi.org/10.1109/tit.2004.834793).
- (2006). "Just relax: convex programming methods for identifying sparse signals in noise". In: *IEEE Transactions on Information Theory* 52.3, pp. 1030–1051. DOI: [10.1109/tit.2005.864420](https://doi.org/10.1109/tit.2005.864420).
- Tuncer, M. A. C. and A. C. Gurbuz (2012). "Ground Reflection Removal in Compressive Sensing Ground Penetrating Radars". In: *IEEE Geoscience and Remote Sensing Letters* 9.1, pp. 23–27. DOI: [10.1109/lgrs.2011.2158981](https://doi.org/10.1109/lgrs.2011.2158981).
- Unde, A. S. and P. Deepthi (2019). "Adaptive Compressive Video Coding for Embedded Camera Sensors: Compressed Domain Motion and Measurements Estimation". In: *IEEE Transactions on Mobile Computing*. DOI: [10.1109/tmc.2019.2926271](https://doi.org/10.1109/tmc.2019.2926271).
- Unser, M. (2000). "Sampling-50 years after Shannon". In: *Proceedings of the IEEE* 88.4, pp. 569–587. DOI: [10.1109/5.843002](https://doi.org/10.1109/5.843002).
- Upadhyay, N. and R. K. Jaiswal (2016). "Single Channel Speech Enhancement: Using Wiener Filtering with Recursive Noise Estimation". In: *Procedia Computer Science* 84, pp. 22–30. DOI: [10.1016/j.procs.2016.04.061](https://doi.org/10.1016/j.procs.2016.04.061).
- Uwaechia, A. N. and N. M. Mahyuddin (2019). "Spectrum-Efficient Distributed Compressed Sensing Based Channel Estimation for OFDM Systems Over Doubly Selective Channels". In: *IEEE Access* 7, pp. 35072–35088. DOI: [10.1109/ACCESS.2019.2904596](https://doi.org/10.1109/ACCESS.2019.2904596).
- Varga, A. and H. J. Steeneken (1993). "Assessment for automatic speech recognition: II. NOISEX-92: A database and an experiment to study the effect of additive noise on speech recognition systems". In: *Speech Communication* 12.3, pp. 247–251. DOI: [10.1016/0167-6393\(93\)90095-3](https://doi.org/10.1016/0167-6393(93)90095-3).
- Vaseghi, S. V. (2005). *Advanced Digital Signal Processing and Noise Reduction*. West Sussex, UK: John Wiley & Sons Ltd.
- Veeraraghavan, A., D. Reddy, and R. Raskar (2010). "Coded strobing photography: Compressive sensing of high speed periodic videos". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33.4, pp. 671–686. DOI: [10.1109/tpami.2010.87](https://doi.org/10.1109/tpami.2010.87).
- Vem, A. et al. (2019). "A User-Independent Successive Interference Cancellation Based Coding Scheme for the Unsourced Random Access Gaussian Channel". In: *IEEE Transactions on Communications* 67.12, pp. 8258–8272. DOI: [10.1109/tcomm.2019.2940216](https://doi.org/10.1109/tcomm.2019.2940216).

- Verhelst, M. and A. Bahai (2015). "Where Analog Meets Digital: Analog-to-Information Conversion and Beyond". In: *IEEE Solid-State Circuits Magazine* 7.3, pp. 67–80. DOI: [10.1109/MSSC.2015.2442394](https://doi.org/10.1109/MSSC.2015.2442394).
- Verteleckaya, E. and K. Sakhnov (2010). "Voice Activity Detection for Speech Enhancement Applications". In: *Acta Polytechnica* 50.4, pp. 100–105.
- Virag, N. (1999). "Single channel speech enhancement based on masking properties of the human auditory system". In: *IEEE Transactions on Speech and Audio Processing* 7.2, pp. 126–137. DOI: [10.1109/89.748118](https://doi.org/10.1109/89.748118).
- Virtanen, T. (2007). "Monaural Sound Source Separation by Nonnegative Matrix Factorization With Temporal Continuity and Sparseness Criteria". In: *IEEE Transactions on Audio, Speech and Language Processing* 15.3, pp. 1066–1074. DOI: [10.1109/tas1.2006.885253](https://doi.org/10.1109/tas1.2006.885253).
- Wadhwa, A., U. Madhow, and N. R. Shanbhag (2017). "Slicer Architectures for Analog-to-Information Conversion in Channel Equalizers". In: *IEEE Transactions on Communications* 65.3, pp. 1234–1246. DOI: [10.1109/TCOMM.2016.2641445](https://doi.org/10.1109/TCOMM.2016.2641445).
- Wang, J. et al. (2016). "Compressive sensing-based speech enhancement". In: *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 24.11, pp. 2122–2131. DOI: [10.1109/TASLP.2016.2598306](https://doi.org/10.1109/TASLP.2016.2598306).
- Wang, J. et al. (2019). "Compressed Sensing Based Selective Encryption With Data Hiding Capability". In: *IEEE Transactions on Industrial Informatics* 15.12, pp. 6560–6571. DOI: [10.1109/tii.2019.2924083](https://doi.org/10.1109/tii.2019.2924083).
- Wang, L., D. Wang, and C. Hao (2019). "A Multiple-Measurement Vectors Reconstruction Method for Low SNR Scenarios". In: *IEEE Transactions on Circuits and Systems II: Express Briefs*. in press, pp. 1–5. DOI: [10.1109/TCSII.2019.2922172](https://doi.org/10.1109/TCSII.2019.2922172).
- Wang, W. et al. (2016). "Combinatorial optimisation for pulse position modulation-ultra wideband signal detection based on compressed sensing and analogue-to-information converter". In: *IET Signal Processing* 10.5, pp. 543–548. DOI: [10.1049/iet-spr.2015.0261](https://doi.org/10.1049/iet-spr.2015.0261).
- Wang, X. et al. (2017). "Look-Ahead Hybrid Matching Pursuit for Multipolarization Through-Wall Radar Imaging". In: *IEEE Transactions on Geoscience and Remote Sensing* 55.7, pp. 4072–4081. DOI: [10.1109/tgrs.2017.2687478](https://doi.org/10.1109/tgrs.2017.2687478).
- Wang, X. et al. (2018). "Two-Level Block Matching Pursuit for Polarimetric Through-Wall Radar Imaging". In: *IEEE Transactions on Geoscience and Remote Sensing* 56.3, pp. 1533–1545. DOI: [10.1109/tgrs.2017.2764920](https://doi.org/10.1109/tgrs.2017.2764920).
- Wang, Y., A. Narayanan, and D. Wang (2014). "On Training Targets for Supervised Speech Separation". In: *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 22.12, pp. 1849–1858. DOI: [10.1109/taslp.2014.2352935](https://doi.org/10.1109/taslp.2014.2352935).
- Wei, H.-T. et al. (2018). *Link Delay Estimation Using Sparse Recovery for Dynamic Network Tomography*. arXiv: [1812.00369](https://arxiv.org/abs/1812.00369) [cs.NI].
- Wei, Z. et al. (2019). "Improving the Signal-to-Noise Ratio of Superresolution Imaging Based on Single-Pixel Camera". In: *IEEE Photonics Journal* 11.1, pp. 1–16. DOI: [10.1109/jphot.2019.2891061](https://doi.org/10.1109/jphot.2019.2891061).
- Wen, F., Y. Tao, and G. Zhang (2015). "Analogue-to-information conversion using multi-comparator-based integrate-and-fire sampler". In: *Electronics Letters* 51.3, pp. 246–247. DOI: [10.1049/el.2014.1950](https://doi.org/10.1049/el.2014.1950).
- Whittaker, E. T. (1915). "XVIII.—On the Functions which are represented by the Expansions of the Interpolation-Theory". In: *Proceedings of the Royal Society of Edinburgh* 35, pp. 181–194. DOI: [10.1017/s0370164600017806](https://doi.org/10.1017/s0370164600017806).
- Wright, J. et al. (2010). "Sparse representation for computer vision and pattern recognition". In: *Proceedings of the IEEE* 98.6, pp. 1031–1044. DOI: [10.1109/JPROC.2010.2044470](https://doi.org/10.1109/JPROC.2010.2044470).

- Wu, D., W. Zhu, and M. N. Swamy (2014). "The Theory of Compressive Sensing Matching Pursuit Considering Time-domain Noise with Application to Speech Enhancement". In: *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 22.3, pp. 682–696. DOI: [10.1109/taslp.2014.2300336](https://doi.org/10.1109/taslp.2014.2300336).
- Wu, D., W. Zhu, and M. N. S. Swamy (2013). "Compressive sensing-based speech enhancement in non-sparse noisy environments". In: *IET Signal Processing* 7.5, pp. 450–457. DOI: [10.1049/iet-spr.2012.0192](https://doi.org/10.1049/iet-spr.2012.0192).
- Wu, D., W.-P. Zhu, and M. Swamy (2011). "A compressive sensing method for noise reduction of speech and audio signals". In: *2011 IEEE 54th International Midwest Symposium on Circuits and Systems (MWSCAS)* (Seoul, South Korea). IEEE, pp. 1–4. DOI: [10.1109/mwscas.2011.6026662](https://doi.org/10.1109/mwscas.2011.6026662).
- Wu, Q. et al. (2015). "Compressive-Sensing-Based High-Resolution Polarimetric Through-the-Wall Radar Imaging Exploiting Target Characteristics". In: *IEEE Antennas and Wireless Propagation Letters* 14, pp. 1043–1047. DOI: [10.1109/lawp.2014.2380787](https://doi.org/10.1109/lawp.2014.2380787).
- Wu, S. et al. (2018). "Single-Pixel Camera in the Visible Band With Fiber Signal Collection". In: *IEEE Access* 6, pp. 17768–17775. DOI: [10.1109/access.2018.2819358](https://doi.org/10.1109/access.2018.2819358).
- Wunder, G. et al. (2015). "Sparse Signal Processing Concepts for Efficient 5G System Design". In: *IEEE Access* 3, pp. 195–208. DOI: [10.1109/access.2015.2407194](https://doi.org/10.1109/access.2015.2407194).
- Xiong, C. et al. (2019). "A Compressed Sensing-Based Element Failure Diagnosis Method for Phased Array Antenna During Beam Steering". In: *IEEE Antennas and Wireless Propagation Letters* 18.9, pp. 1756–1760. DOI: [10.1109/LAWP.2019.2929353](https://doi.org/10.1109/LAWP.2019.2929353).
- Xiong, T. et al. (2018). "An Unsupervised Compressed Sensing Algorithm for Multi-Channel Neural Recording and Spike Sorting". In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 26.6, pp. 1121–1130. DOI: [10.1109/tnsre.2018.2830354](https://doi.org/10.1109/tnsre.2018.2830354).
- Xu, H., C. Zhang, and I. Kim (2019). "Coupled Online Robust Learning of Observation and Dictionary for Adaptive Analog-to-Information Conversion". In: *IEEE Signal Processing Letters* 26.1, pp. 139–143. DOI: [10.1109/LSP.2018.2880566](https://doi.org/10.1109/LSP.2018.2880566).
- Xu, L. et al. (2018). "Joint Two-Dimensional DOA and Frequency Estimation for L-Shaped Array via Compressed Sensing PARAFAC Method". In: *IEEE Access* 6, pp. 37204–37213. DOI: [10.1109/ACCESS.2018.2850307](https://doi.org/10.1109/ACCESS.2018.2850307).
- Xu, P. et al. (2007). "Lp Norm Iterative Sparse Solution for EEG Source Localization". In: *IEEE Transactions on Biomedical Engineering* 54.3, pp. 400–409. DOI: [10.1109/tbme.2006.886640](https://doi.org/10.1109/tbme.2006.886640).
- Xu, W. et al. (2019). "An Efficient Wideband Spectrum Sensing Algorithm for Unmanned Aerial Vehicle Communication Networks". In: *IEEE Internet of Things Journal* 6.2, pp. 1768–1780. DOI: [10.1109/JIOT.2018.2882532](https://doi.org/10.1109/JIOT.2018.2882532).
- Xu, Y. et al. (2017). "A Survey of Dictionary Learning Algorithms for Face Recognition". In: *IEEE Access* 5, pp. 8502–8514. DOI: [10.1109/ACCESS.2017.2695239](https://doi.org/10.1109/ACCESS.2017.2695239).
- Yang, A. et al. (2013). "Fast ℓ_1 -Minimization Algorithms for Robust Face Recognition". In: *IEEE Transactions on Image Processing* 22.8, pp. 3234–3246. DOI: [10.1109/TIP.2013.2262292](https://doi.org/10.1109/TIP.2013.2262292).
- Yang, J. and Y. Zhang (2011). "Alternating Direction Algorithms for ℓ_1 -Problems in Compressive Sensing". In: *SIAM Journal on Scientific Computing* 33.1, pp. 250–278. DOI: [10.1137/090777761](https://doi.org/10.1137/090777761).
- Yang, X. et al. (2013). "Energy-Efficient Distributed Data Storage for Wireless Sensor Networks Based on Compressed Sensing and Network Coding". In: *IEEE Transactions on Wireless Communications* 12.10, pp. 5087–5099. DOI: [10.1109/TWC.2013.090313.121804](https://doi.org/10.1109/TWC.2013.090313.121804).
- Yang, Y., X. Qin, and B. Wu (2019). "Fast and accurate compressed sensing model in magnetic resonance imaging with median filter and split Bregman method". In: *IET Image Processing* 13.1, pp. 1–8. DOI: [10.1049/iet-ipr.2018.5173](https://doi.org/10.1049/iet-ipr.2018.5173).

- Yin, W. et al. (2008). "Bregman Iterative Algorithms for ℓ_1 -Minimization with Applications to Compressed Sensing". In: *SIAM Journal on Imaging Sciences* 1.1, pp. 143–168. DOI: [10.1137/070703983](https://doi.org/10.1137/070703983).
- Yu, N. Y. (2019). "Multiuser Activity and Data Detection Via Sparsity-blind Greedy Recovery for Uplink Grant-Free NOMA". In: *IEEE Communications Letters* 23.11, pp. 2082–2085. DOI: [10.1109/lcomm.2019.2937117](https://doi.org/10.1109/lcomm.2019.2937117).
- Yuan, L. et al. (2019). "Analysis $L_{1/2}$ Regularization: Iterative Half Thresholding Algorithm for CS-MRI". In: *IEEE Access* 7, pp. 79366–79373. DOI: [10.1109/ACCESS.2019.2923171](https://doi.org/10.1109/ACCESS.2019.2923171).
- Zand, T. et al. (2019). "Compressed Imaging to Reduce Storage in Adjoint-State Calculations". In: *IEEE Transactions on Geoscience and Remote Sensing* 57.11, pp. 9236–9241. DOI: [10.1109/tgrs.2019.2925684](https://doi.org/10.1109/tgrs.2019.2925684).
- Zayyani, H., R. Sari, and M. Korke (2017). "A Distributed 1-bit Compressed Sensing Algorithm for Nonlinear Sensors With a Cramer–Rao Bound". In: *IEEE Communications Letters* 21.12, pp. 2626–2629. DOI: [10.1109/LCOMM.2017.2748943](https://doi.org/10.1109/LCOMM.2017.2748943).
- Zeng, F., C. Li, and Z. Tian (2011). "Distributed Compressive Spectrum Sensing in Cooperative Multihop Cognitive Networks". In: *IEEE Journal of Selected Topics in Signal Processing* 5.1, pp. 37–48. DOI: [10.1109/JSTSP.2010.2055037](https://doi.org/10.1109/JSTSP.2010.2055037).
- Zhang, P. and J. Wang (2019). "On enhancing network dynamic adaptability for compressive sensing in WSNs". In: *IEEE Transactions on Communications* 67.12, pp. 8450–8459. DOI: [10.1109/TCOMM.2019.2938950](https://doi.org/10.1109/TCOMM.2019.2938950).
- Zhang, R., H. Zhao, and J. Zhang (2018). "Distributed Compressed Sensing Aided Sparse Channel Estimation in FDD Massive MIMO System". In: *IEEE Access* 6, pp. 18383–18397. DOI: [10.1109/ACCESS.2018.2818281](https://doi.org/10.1109/ACCESS.2018.2818281).
- Zhang, S., X. Zhao, and B. Lei (2012). "Robust facial expression recognition via compressive sensing". In: *Sensors* 12.3, pp. 3747–3761. DOI: [10.3390/s120303747](https://doi.org/10.3390/s120303747).
- Zhang, W. et al. (2011). "Virtual probe: a statistical framework for low-cost silicon characterization of nanoscale integrated circuits". In: *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 30.12, pp. 1814–1827. DOI: [10.1109/TCAD.2011.2164536](https://doi.org/10.1109/TCAD.2011.2164536).
- Zhang, W. and A. Hoorfar (2015). "A Generalized Approach for SAR and MIMO Radar Imaging of Building Interior Targets With Compressive Sensing". In: *IEEE Antennas and Wireless Propagation Letters* 14, pp. 1052–1055. DOI: [10.1109/lawp.2015.2394746](https://doi.org/10.1109/lawp.2015.2394746).
- Zhang, X. et al. (2018). "Compressive Sensing-Based Multiuser Detection via Iterative Reweighted Approach in M2M Communications". In: *IEEE Wireless Communications Letters* 7.5, pp. 764–767. DOI: [10.1109/LWC.2018.2820704](https://doi.org/10.1109/LWC.2018.2820704).
- Zhang, Y., J. Yang, and W. Yin (2011). *YALL1 Basic: A solver for sparse reconstruction*. <http://ya111.blogs.rice.edu>. v.1.4.
- Zhang, Y. et al. (2016). "A Review of Compressive Sensing in Information Security Field". In: *IEEE Access* 4, pp. 2507–2519. DOI: [10.1109/ACCESS.2016.2569421](https://doi.org/10.1109/ACCESS.2016.2569421).
- Zhang, Y. et al. (2019). "Efficient Estimation and Prediction for Sparse Time-Varying Underwater Acoustic Channels". In: *IEEE Journal of Oceanic Engineering*, pp. 1–14. DOI: [10.1109/JOE.2019.2911446](https://doi.org/10.1109/JOE.2019.2911446).
- Zhang, Y. and M. Xing (2019). "Joint Method of ISAR Imaging and Scaling for Maneuvering Targets via Compressive Sensing". In: *IEEE Sensors Journal* 19.17, pp. 7300–7307. DOI: [10.1109/jsen.2019.2897383](https://doi.org/10.1109/jsen.2019.2897383).
- Zhang, Z. et al. (2019). "Electrocardiogram Reconstruction Based on Compressed Sensing". In: *IEEE Access* 7, pp. 37228–37237. DOI: [10.1109/access.2019.2905000](https://doi.org/10.1109/access.2019.2905000).

- Zhang, Z. and Y. Shen (2019). "Listener Preference on the Local Criterion for Ideal Binary-Masked Speech". In: *the 20th Annual Conference of the International Speech Communication Association (INTER-SPEECH)* (Graz, Austria). ISCA, pp. 1383–1387. DOI: [10.21437/interspeech.2019-1369](https://doi.org/10.21437/interspeech.2019-1369).
- Zhao, W. et al. (2018). "On-Chip Neural Data Compression Based On Compressed Sensing With Sparse Sensing Matrices". In: *IEEE Transactions on Biomedical Circuits and Systems* 12.1, pp. 242–254. DOI: [10.1109/tbcas.2017.2779503](https://doi.org/10.1109/tbcas.2017.2779503).
- Zhao, Y. et al. (2017). "Super Resolution Imaging Based on a Dynamic Single Pixel Camera". In: *IEEE Photonics Journal* 9.2, pp. 1–11. DOI: [10.1109/jphot.2017.2688351](https://doi.org/10.1109/jphot.2017.2688351).
- Zhou, S. et al. (2019). "A Distributed Compressive Data Gathering Framework For Mobile Crowdsensing". In: *IEEE Internet of Things Journal*, pp. 1–1. DOI: [10.1109/JIOT.2019.2921203](https://doi.org/10.1109/JIOT.2019.2921203).
- Zong, Y. et al. (2016). "Double sparse learning model for speech emotion recognition". In: *Electronics Letters* 52.16, pp. 1410–1412. DOI: [10.1049/el.2016.1211](https://doi.org/10.1049/el.2016.1211).
- Zonoobi, D., A. A. Kassim, and Y. V. Venkatesh (2011). "Gini Index as Sparsity Measure for Signal Reconstruction from Compressive Samples". In: *IEEE Journal of Selected Topics in Signal Processing* 5.5, pp. 927–932. DOI: [10.1109/jstsp.2011.2160711](https://doi.org/10.1109/jstsp.2011.2160711).