Royal Military Academy
Polytechnic Faculty
Brussels — Belgium

University of Liège
Faculté des Sciences Appliquées
Liège — Belgium

Space-time adaptive processing for multistatic radars

Xavier Neyt

Supervisors: Prof. Dr. Ir. Marc. Acheroy, Royal Military Academy
Prof. Dr. Ir. Jacques. Verly, University of Liège

Thesis submitted in fulfillment of the requirement for the degree of
Doctor in Engineering Sciences (RMA)
Docteur en Sciences de l'Ingénieur (ULg).

July 2008

Committee

Chairman:

Dr. Marc Lesturgie
Office National d'Etudes et de Recherches Aéronautiques, France

Supervisors:

Prof. Dr. Ir. Marc Acheroy
Royal Military Academy

Prof. Dr. Ir. Jacques Verly
Université de Liège

Members:

Prof. Dr. Hugh Griffiths
Cranfield University, UK

Dr. Michael Wicks
Air Force Research Laboratory, USA

Dr. Christian Barbier
University of Liège

Prof. Dr. Ir. Christiaan Perneel
Royal Military Academy

Public defense held July 3rd 2008.

Acknowledgments

Although the redaction of a thesis is supposed to be a solitary exercise, it cannot be done without the support from many people. First and foremost, I would like to thank Prof. Marc Acheroy, head of the Communication, Information, Systems and Sensors department of the Royal Military Academy and founder of the Signal and Image Centre, for the freedom and the support he gave me. I hope I made a good use of that freedom. I am also in debt to Col. Jean Marsia, Director of Academic Affairs, without whom the Royal Military Academy would not be able to confer Ph.D. degrees. His action were also the key to get the cotutelle between RMA and the University of Liège in place. Of course, I also wish to thank Prof. Jacques Verly, my supervisor at the University of Liège. He introduced me to Space-Time Adaptive Processing during his memorable “Francqui” courses held at the Université Libre de Bruxelles and we often had lengthy exchanges about subtle technical issues. As far as radars are concerned, I have to thank Pascal Lecomte, Head of Data Quality and Algorithms at the European Space Agency. You get to learn a lot of subtle things when you have to calibrate a radar system to less than 0.01dB. Similarly, I think that it is Prof. Helmut Süß, from the German Aerospace Center, that introduced me to the random scattering of clutter and the various clutter models. I would also like to thank Michael Wicks from the US Air Force Laboratory and Prof. Hugh Griffiths from the Cranfield University not only for accepting to evaluate this thesis but also for the discussion we had around a glass of Belgian beer.

Regarding the subject of the thesis, I am deeply in debt to Fabian Lapierre, who guided my first steps in STAP when he was himself finishing his Ph.D. at the University of Liège. I would also like to thank Pauline Pettiaux for her indestructible optimism. We worked together on the ERS scatterometer processor and she might have been the first to see some STAP results. For the measurements of Chapter 6, I have to thank Mireille and Virginie Kubica, who did their master thesis on the use of transmitters of opportunity, for their incredible dedication, hard work, and faith, and Cdt. Jacques Raout from the French Air Force Academy for providing them (and me) the opportunity to work with him. Key support for setting up the contact with the French Air Force Academy was provided by Col. Patrice Laurent. I wish also to thank Philippe Ries, Ph.D. student at the University of Liège, for his contribution to Chapter 5.

Finally, I’d also like to thank the colleagues at the Signal and Image Centre. In particular Pascal Druyts for his contribution to Chapter 7 and more generally for his deep and profound insight into the most diverse subjects. This often lead to extremely lengthy discussions about technical (and, sometimes, non-technical) subjects. I also wish to thank Peggy Decroix not only for taking over my duties whenever I was unavailable but also for the puzzles and the chats. And last but certainly not least, I’d wish to express my gratitude to Nathalie Pattyn for her support, her contagious dynamism, her indefectible trust and her friendship.

Thank you all.

Contents

Chapter 1

Notations

1.1 Symbols

$\sim$	Is distributed according to.
$\circ$	Hadamard product (element-wise).
$\otimes$	Kronecker product [34, Chapter 2].
$\vec{1}$	A unit vector.
$\vec{1}_x$	A unit vector parallel to the x-axis.
$\vec{1}_{RS}$	A unit vector from the receiver to the scatterer.
$\vec{1}_{TS}$	A unit vector from the transmitter to the scatterer.
$\{\mathbf{A}\}_d$	A vector containing the diagonal elements of the matrix $\mathbf{A}$.
α	Complex amplitude of the target signal.
$\boldsymbol{\alpha}$	Vector of complex amplitude of the scatterers along an isorange.
α_p	Complex amplitude of the target signal received by radar p.
α_R	Angle between the velocity vector and the x-axis.
$\mathbf{a}$	Spatial steering vector.
$A_{p,k}$	Area of clutter patch k of radar p.
$\mathbf{b}$	Temporal steering vector.
B	Spatio-temporal beam pattern associated with the signal match power spectrum estimator.
B_{MVE}	Spatio-temporal beam pattern associated with the MVE power spectrum estimator.
c	speed of light.
$c_p(\vec{x})$	Amplitude factors for radar p and a scatterer located at $\vec{x}$.
$\mathbb{C}$	The set of all complex numbers.
$\mathrm{diag}\{\mathbf{a}\}$	A diagonal matrix with vector $\mathbf{a}$ on the diagonal.
$\mathbf{D}$	Normalization matrix for the matched filter estimator.
$\eta_{b,m}$	Decision threshold in the bistatic case when a Marcum target model is considered.
$\eta_{b,s}$	Decision threshold in the bistatic case when a Swerling-I target model is considered.

$\eta_{m,m}$	Decision threshold in the multistatic case when a Marcum target model is considered.
$\eta_{m,s}$	Decision threshold in the multistatic case when a Swerling-I target model is considered.
$E(a)$	Expectation of the random variable a.
$E(a\|b)$	Conditional mean of the random variable a conditioned by the random variable b.
$E(a;b)$	Expectation of the random variable a with PDF parameterized by the deterministic value b.
$\phi(\vec{r})$	Phase shift at the antenna element located at $\vec{r}$.
ϕ_T	Azimuth of the receiver w.r.t. the transmitter.
ϕ_R	Azimuth of the receiver w.r.t. the receiver.
f_c	Frequency of the carrier.
f_{D_R}	Doppler frequency due to the relative movement between the receiver and the scatterer.
f_{D_T}	Doppler frequency due to the relative movement between the transmitter and the scatterer.
f_p	Sampling frequency (PRF).
F	Cumulative density function.
$\mathbf{F}$	Estimation filter.
γ_0	Gamma naught.
H	Height of the transmitter above the ground.
H_0	Null hypothesis (no target).
H_1	Alternative hypothesis (target present).
$\mathbf{I}_M$	Identity matrix of size $M \times M$. The subscript denoting the size can be dropped if it is obvious from the context.
$\vec{k}$	Wave vector.
K	Number of clutter patches along an isorange.
$L(H_0)$	likelihood of hypothesis H_0.
$l(H_0)$	log-likelihood of hypothesis H_0.
λ	Carrier wavelength.
λ_c	Carrier wavelength.
$\Lambda(\mathbf{y})$	likelihood ratio $\frac{L(H_1)}{L(H_0)}$.
$\lambda(\mathbf{y})$	log-likelihood ratio $\ln \frac{L(H_1)}{L(H_0)}$.
$\Lambda_g(\mathbf{y})$	generalized likelihood ratio.
$\lambda_g(\mathbf{y})$	generalized log-likelihood ratio.
M	Number of pulses in a CPI.
μ	Non-centrality parameter.
ν_D	Reduced Doppler frequency.

ν_{p_D}	Reduced Doppler frequency corresponding to radar p.
$\vec{\nu}_s$	Reduced spatial frequency vector ($\vec{\nu}_s = (\nu_{s,x}, \nu_{s,y}, \nu_{s,z})$).
$\vec{\nu}_{p_s}$	Reduced spatial frequency vector corresponding to radar p.
$\mathbf{n}$	Thermal noise vector.
N	Number of antenna receive channels.
N_p	Number of antenna receive channels for radar p.
N_r	Number of range rings.
ψ	Ambiguity function.
ψ_g	Generalized ambiguity function.
$p(a)$	PDF of the random variable a.
$p(a\|b)$	PDF of the random variable a conditioned by the random variable b.
$p(a;b)$	PDF of the random variable a parametrized by the deterministic parameter b.
$p(t)$	Complex amplitude of the transmitted signal.
$\mathbf{p}$	Vector of samples of the complex amplitude of the transmitted signal.
P_{FA}	Probability of false alarm.
P_{D}	Probability of detection.
$P(\vec{k}, \omega)$	Power spectrum.
P_{SM}	Signal match power spectrum estimator.
P_{PER}	Periodogram-based power spectrum estimator.
P_{MVE}	Minimum variance power spectrum estimator.
Q	Complement cumulative density function.
$\vec{\rho}_{p,n}$	Reduced location of the n-th antenna receive element of radar p.
$r(\vec{\Delta r}, \Delta t)$	Covariance function.
$\mathbf{r}_\alpha$	Diagonal elements of the covariance matrix of $\boldsymbol{\alpha}$.
$\mathbf{R}$,$\mathbf{R}_{i+n}$	Interference plus noise covariance matrix.
$\mathbf{R}_c$	Clutter covariance matrix.
$\mathbf{R}_p$	Interference plus noise covariance matrix corresponding to bistatic radar p.
$\mathbf{R}_\alpha$	Covariance matrix of $\boldsymbol{\alpha}$.
$\tilde{\mathbf{R}}_\alpha$	*A priori* covariance matrix of $\boldsymbol{\alpha}$.
$\hat{\mathbf{R}}_p$	Estimated interference plus noise covariance matrix corresponding to bistatic radar p.
$\mathbf{R}_{c_p}$	Clutter covariance matrix corresponding to bistatic radar p.
R	Bistatic range.
R_{RT}	Distance between the transmitter and the receiver.
R_R	Distance between the scatterer and the receiver.
R_T	Distance between the scatterer and the transmitter.
σ_0	Sigma naught.
$\sigma_{0_{p_k}}$	Sigma naught of clutter patch k for radar p.

σ_α^2	Variance of α.
σ_n^2	Variance of the (thermal) noise.
$\mathbf{s}$	Steering vector (including amplitude factor).
$\mathbf{s}_p$	Steering vector corresponding to radar p (including amplitude factor).
S	Subsampling factor.
$\mathbf{S}$	Matrix of steering vectors (including amplitude factors).
SNR_p	Signal to noise ratio corresponding to radar p.
SINR_p	Signal to interference-plus-noise ratio corresponding to radar p.
$\mathrm{SINR}_{\mathrm{loss}p}$	Signal to interference-plus-noise ratio loss corresponding to radar p.
θ	Elevation of the receiver w.r.t. the transmitter.
θ_T	Elevation of the scatterer w.r.t. the transmitter.
θ_R	Elevation of the scatterer w.r.t. the receiver.
$\boldsymbol{\theta}$	Parameter vector.
$\boldsymbol{\theta}_p$	Parameter vector corresponding to radar p.
$\tau'(\vec{r})$	Additional delay between the phase center and the antenna element located at $\vec{r}$.
τ_{rt}	Round-trip time from the transmitter to the receiver antenna's phase center.
T_c	Period of the carrier.
T_{CPI}	Duration of the CPI.
$T_{b,m}$	Test statistic in the bistatic case considering a Marcum target model.
$T_{b,s}$	Test statistic in the bistatic case considering a Swerling-I target model.
$T_{m,m}$	Test statistic in the multistatic case considering a Marcum target model.
$T_{m,s}$	Test statistic in the multistatic case considering a Swerling-I target model.
$\vec{u}$	Unit vector anti-parallel to $\vec{k}$.
$\vec{v}_R$	Velocity vector of the receiver.
v_R	Magnitude of the velocity vector of the receiver.
$\vec{v}_S$	Velocity vector of the scatterer.
v_S	Magnitude of the velocity vector of the scatterer.
$\vec{v}_T$	Velocity vector of the transmitter.
v_T	Magnitude of the velocity vector of the transmitter.
$\mathbf{v}$	Steering vector.
$\mathbf{v}_p$	Steering vector corresponding to radar p.
χ_ν^2	Central chi-squared distribution with ν degrees of freedom.
$\chi_\nu'^2(\lambda)$	Non-central chi-squared distribution with ν degrees of freedom and non-centrality parameter λ.
$\mathbf{W}$	Spatial filtering matrix.
$\mathbf{w}$	Optimum detection filter.
$\mathbf{w}_p$	Optimum detection filter corresponding to radar p.
$\hat{\mathbf{w}}_p$	Sub-optimum detection filter corresponding to radar p.

$\vec{x}$	A geometric vector; the position of a scatterer.
$\mathbf{x}$	A vector.
$\mathbf{x}^*$	The complex conjugate of vector $\mathbf{x}$ (not transposed). The same notation is also applied to matrices.
$\mathbf{x}^\dagger$	The transposed conjugate of vector $\mathbf{x}$. The same notation is also applied to matrices.
$\|\mathbf{x}\|$	The l^2 norm of vector $\mathbf{x}$.
$\mathbf{X}$	A matrix.
$\lvert\mathbf{X}\rvert$	The determinant of matrix $\mathbf{X}$.
$y_p(n,m)$	Measurement sample corresponding to the m-th pulse and obtained at the n-th antenna element of radar p.
$\mathbf{y}$	A spatio-temporal measurement vector.
$\mathbf{y}_p$	The spatio-temporal measurement vector corresponding to bistatic radar p.
ζ	Angle between the velocity vector of the scatterer and the x-axis.

1.2 Abbreviations

AWACS	Airborne Warning and Control System.
CDF	Cumulative Density Function.
CMF	Crude Matched Filter.
CPI	Coherent Processing Interval.
CFAR	Constant False Alarm Rate.
DDC	Digital Down Conversion.
DVB-T	Digital Video Broadcast - Terrestrial.
EM	Expectation-Maximization.
ECM	Electronic Counter Measures.
ECCM	Electronic Counter Counter Measures.
FIR	Finite Impulse Response.
GLRT	Generalized Likelihood Ratio Test.
GMF	Generalized Matched Filter.
GSM	Global System for Mobile communications.
ICM	Internal Clutter Motion.
LNA	Low Noise Amplifier.
LRT	Likelihood Ratio Test.
MAP	Maximum a Posteriori.
MF	Matched filter.
MIMO	Multiple Inputs, Multiple Outputs.
MLE	Maximum Likelihood Estimate.
MSMI	Modified Sample Matrix Inversion.
MVE	Minimum Variance Estimate (or Estimator).
PDF	Probability Density Function.

PRF	Pulse Repetition Frequency.
PRI	Pulse Repetition Interval.
RCS	Radar Cross Section.
RF	Radio Frequency.
Rx	Receiver.
SCM	Sample Covariance Matrix.
SINR	Signal to Interference plus Noise Ratio.
SMI	Sample Matrix Inversion.
STAP	Space-Time Adaptive Processing.
TMF	Tapered matched filter.
Tx	Transmitter.
UAV	Unmanned Aerial Vehicle.
ULA	Uniform Linear Array.

Chapter 2

Introduction

This work addresses the detection of ground moving targets using multistatic radars. This chapter puts the studied concepts in perspective and concludes with the general structure of the thesis

2.1 Ground moving target detection

The detection of vehicles moving on the ground is a major challenge in airborne and space borne radar. Generally speaking, the detection performance of a radar is limited by the signals that compete with the target return. When the radar receives no return from the ground, e.g. as it is looking at air targets, detection is limited by the thermal noise, typically white and of low amplitude. On the other hand, ground target echoes always compete with clutter echoes. Clutter echoes are caused by objects such as buildings or vegetation that are not interesting in our context and are typically not moving. Hence, targets contributions can mostly be distinguished from clutter based on the Doppler frequency shift of the received echoes. This is illustrated in Figure 2.1 where the — idealized — spectrum of a signal received by a non-moving radar is depicted. The clutter component is clearly visible at a Doppler frequency of 0Hz and a target

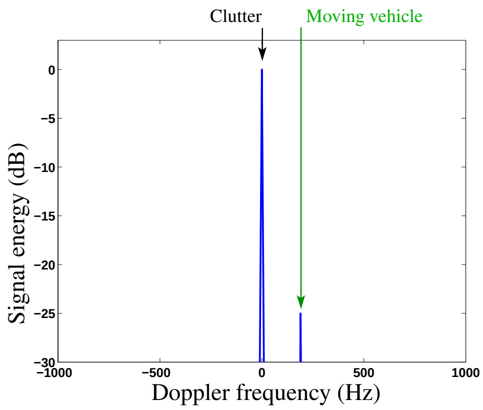
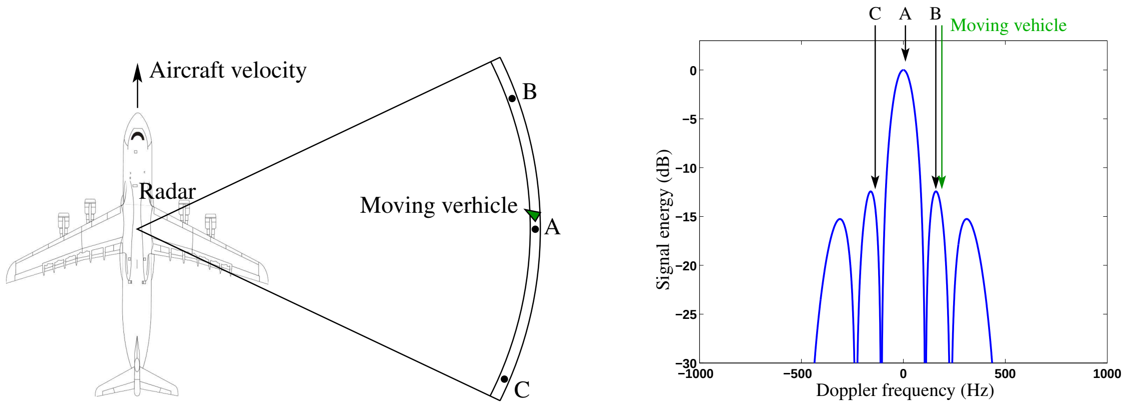


Figure 2.1: Spectrum of the received signal with a non-moving radar.

component around 200Hz.

With a fixed radar, the spectral behavior of the clutter is relatively easy to model and the filter to remove it can, in first approximation, be kept constant and independent of the actual scenario. However, if the radar system is moving, the clutter spectral signature can be extremely complex.



(a) (b)

Figure 2.2: (a) Side looking airborne radar and (b) spectrum of the received signal.

A scenario involving an airborne radar is illustrated in Figure 2.2 (a), and the spectrum of the received signal is illustrated in Figure 2.2 (b). Obviously, the signal reflected from location A exhibits a Doppler frequency shift of 0Hz as point A has no radial relative velocity component. Since the radar is approaching B, the frequency shift is positive for the signal reflected from location B whereas it is negative for the signal reflected from location C. In this scenario, it turns out that the weak moving target signal is hidden by the clutter signal from location B. Hence, the Doppler frequency dimension alone does not permit a separation of both signals.

Separation of the clutter echoes from target echoes requires the introduction of additional information. If the receiver is equipped with a receiver array, the signals can also be distinguished

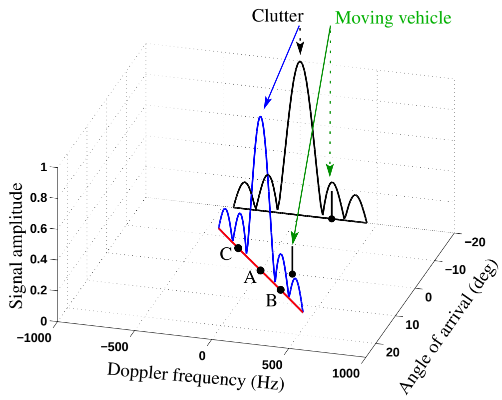


Figure 2.3: Spatio-temporal representation of the received signal.

by their direction of arrival, i.e., by adding a spatial information. In the previous example, the signals from locations A, B and C have indeed different directions of arrival. To illustrate this, Figure 2.3 depicts an angle-Doppler representation of the signal. The blue curve is the (idealized) spatio-temporal spectrum of the signal. The black curve is the projection of the blue one and is the same as the one depicted in Figure 2.2 (b), i.e., a Doppler frequency representation. The angle-Doppler location of the points A, B and C is indicated on the graph. If the target is moving, although it has the same Doppler signature as signals coming from location B, it has not the same direction of arrival. This illustrates how additional spatial information permits the separation of targets from clutter even with a moving radar.

Further, a filter is required to cancel or attenuate the signal reflected by clutter. The computation of this filter is not straightforward. Moreover, as the scenario is inherently dynamic, adaptation to the changing environment is required in order to secure the same detection performance. The computation of this adaptive filter is a central issue in space-time adaptive processing (STAP).

2.2 Passive radars

Most current radars are monostatic, i.e., the receiver and the transmitter are collocated, which simplifies the synchronization of the transmitter and the receiver. However, the emission of the radiation also discloses the location of the radar.

In bistatic radars, the receiver and the transmitter are separated, and sometimes located very far from each other. Initially, the motivation was purely technical as a good isolation between the transmitter and the receiver is required for proper operation which was difficult to achieve in the past. Current motivations [107] include an increased sensitivity thanks to possible longer integration time, less sensitivity to target's stealthiness, and low probability of intercept as the receiver platform is totally passive.

Particularly challenging implementations of bistatic radars are the passive radars, typically characterized by the use of transmitters of opportunity. Whereas this type of radar is certainly not new [58, 179], its first implementations were plagued by the poor characteristics of the available signals of opportunity. This type of radar is regaining some interest with the advent of signal sources with digital modulation, that have more favorable characteristics [56, 147].

2.3 Multistatic radars

Multistatic [14, 15, 18, 20, 24, 36, 65] radars are characterized by the presence of several transmitters and/or several receivers, where the receivers cooperate in order to detect and locate targets. This type of radar is also referred to as Multiple Input Multiple Output (MIMO) radars [41, 141, 148] in analogy with MIMO communication channels, where the transmitter and the receiver system corresponds to several transmitting and receiving sites. This type of radar is also referred to as multilateral radar [76] when no coherence between the receivers is assumed, netted radar or network of radars [10, 135], and distributed array [3, 4, 143] when the receivers are assumed coherent among each other. Particular cases of multistatic radars are:

A group of bistatic radars, where the output of the bistatic radars are processed centrally to obtain a decision regarding the presence of a target and to estimate its parameters (speed, radar cross section). In this context, it is assumed that the transmitters do not interfere with each other, which is typically achieved either by using separate frequency bands or orthogonal transmitted waveforms [48, 77]. Further, each receiver is assumed to be able to receive the signals from each transmitter.

A single transmitter with several receivers, typically in the case of a high value unit equipped with the transmitter, for instance an AWACS, and receivers cooperating to achieve the detection.

A single receiver with several transmitters, for instance in the case of a single receiver exploiting several transmitters of opportunity operating in different frequency bands.

These radars can cooperate coherently or not. If the transmitters are coherent, they can operate as a large thinned array and do beamforming on transmit. However, as the baseline between the different transmitters is mostly much larger than the wavelength, this beamforming results in poor sidelobes. Similarly, the receiver can operate coherently and form a large thinned array to do beamforming on receive. Obviously, coherency of the local oscillator among different receivers (transmitters) is not a sufficient condition to permit coherent operation, as the scatterer also needs to be coherent among the different transmitter/receivers, which is only the case in particular situations. In the case of angle-only multistatic radars, which do not have any range resolution, target localization is still possible by exploiting the spatial (angular) resolution of each radar [143, 144].

As far as performance is concerned, a multistatic configuration takes advantage of

- the geometric gain, as, under favorable geometry, the angular diversity can indeed be exploited to cancel the phenomenon of blind speed [122];
- the angular diversity, as the detection probability can be enhanced by taking advantage of the decorrelation arising from the different look angles [42];
- the anti-stealth properties, as a stealth target often minimizes the monostatic backscatter by scattering the energy in some other direction than the direction of arrival.

As far as tactical considerations are concerned, a multistatic configuration benefits from

- stealth operations, as the transmitter can be placed at a safe stand-off distance while the receiver(s) are totally silent and hence difficult to detect;
- the redundancy, as well on the receiving side as on the transmitting side. The detection is still possible, however with decreased performance, when one device (receiver, transmitter) fails.

- the deception, as several transmitters can be used either together or successively in order to deceive counter measures. A passive (opportunistic) radar using existing radio frequency (RF) sources is an interesting particular case.

- anti-ECM, in particular anti-jamming (directional jammer or intelligent repeating jammer) in the surrounding of the target as this kind of counter-measure is not effective without knowledge of the receiver location.

As far as technologies are concerned, a multistatic configuration

- allows for continuous wave operation. As the transmitter does not need to be silent when the receiver is operating, the radar can exploit continuous waves and hence increase the transmitted energy while preserving a low probability of intercept;

- avoid the use of a duplexer, as it is not necessary to switch between a transmit and a receive mode, which simplifies the construction of the radar;

- increases mobility, as the transmitter and the receiver are distinct devices, they can be operated in different conditions. The performance of a transmitter (average transmit power, duration of operation) is typically limited by the available electrical power. The power requirement in turn limits the mobility of the transmitter. For instance, it is well known that the available power during the eclipse part of the orbit, i.e., when the spacecraft is operating on its batteries, is limiting the operation of space-borne radars. During the non-eclipse part of the orbit, limited heat transfer (cooling) is one of the factors that limits the performance. On the contrary, a receiver can be very small and consumes little power.

On the other hand, multistatic radars have drawbacks. As far as performance is concerned, the bistatic scattered energy and hence the bistatic radar cross section is smaller than in the case of a monostatic radar configuration due to the removal of the specular returns from corner reflectors and to shadowing effects.

From a tactical point of view,

- the practical positioning of receivers and transmitters is critical to avoid shadowing effects;

- the operation and the maintenance of several units at different locations can become complex.

From a technical point of view,

- the synchronization of the transmitter and the receiver can be problematic except in the case of a free line of sight or of a direct link between the transmitter and the receiver;

- the temporal synchronization of the receivers in case of coherent operations, can be complex;

- the spatial synchronization (pulse chasing) can be challenging;

- in case of a centralized processing, the data necessary to perform the detection needs to be transmitted to the control center, which implies high bandwidth links. A mitigation of this problem consists in performing a decentralized detection, however with degraded performance.

2.4 Thesis structure

Chapter 3 details the basic geometric relationships and the signal models.

Chapter 4 presents the target detection theory together with performance measures used throughout this thesis. Moreover it extends the target detection theory to multistatic detections in the presence of colored noise in a systematic and homogeneous way. Finally, it assesses the actual usefulness of multistatism.

Chapter 5 develops a detailed modeling of the power spectrum of the clutter signal. The classical angle-Doppler presentation is shown to be the projection of the more general four-dimensional power spectrum.

Chapter 6 analyzes arbitrary waveforms as radar waveforms and focuses on the feasibility of using transmitters of opportunity as signal source for space-time processing. The synchronization — or lack thereof — is a crucial issue in the case of transmitters of opportunity. The proposed approach immediately extends to multistatic systems.

Chapter 7 focuses on the estimation of the clutter covariance matrix in a non-homogeneous range-dependent environment. The proposed method, based on the maximum a posteriori estimate of the ground reflectivity, uses the radar measurements themselves. Finally, this chapter considers possible extensions to multistatic systems.

Finally, the concluding Chapter 8 gives some perspectives for further research.

2.5 Original contributions

Our main original contributions are

- A rigorous and uniform analysis of multistatic target detection in the presence of colored noise is provided. Although similar results appear in the literature, these are typically limited to white noise or known target echo amplitude.

- The 4D clutter power spectrum locus is introduced. This greatly simplifies the analysis of the clutter power spectrum behavior as a function of radar configuration parameters. In particular, this permits us to provide a definitive rule about the configurations leading to a range-independent clutter covariance matrix. This also provides a means to study the influence of different antenna configurations, i.e., non-ULA antennas.

- Space-time adaptive processing is extended to arbitrary waveforms. This generalizes the train of coherent pulses usually considered. In addition, we provide a means to perform the STAP efficiently, i.e., without requiring the inversion of a gigantic matrix. We note that arbitrary waveforms are commonly used in passive radars, but to our knowledge, the combined use of the spatial and the temporal dimension was never considered.

- A method to estimate the covariance matrix, taking into account its particular structure, is proposed. Using this method, a justification is provided for the intuitively appealing spatial averaging usually performed.

Chapter 3

Signal modeling

3.1 Introduction

This chapter provides a basic introduction to spatio-temporal modeling of the signal with as goal to obtain the basic relations used in the subsequent chapters. A relatively simple model is deliberately considered in order not to obscure fundamental issues. A more complex model, which accounts for polarization and element-dependent antenna radiation pattern is found in [70].

The considered multistatic radar consists of several non collocated mutually coherent bistatic radars. It is assumed that each receiver can receive the signal from each transmitter and the waveforms used by the different transmitters are assumed not to interfere. The latter assumption is verified for instance when each transmitter operates in a different frequency band or when the transmitter uses techniques analogous to Code Division Multiple Access communication [48]. Furthermore, the considered transmitters and receivers are moving platforms, for instance Unmanned Aerial Vehicles. The receivers carry an antenna array of arbitrary shape with several channels.

The relations for a bistatic radar are first developed and then extend to several bistatic radars, i.e., a multistatic radar.

3.2 Geometry

3.2.1 Bistatic transmitter/receiver geometric configuration

The considered bistatic geometric configuration for an arbitrary array is illustrated in Figure 3.1 and defines the positions and velocities of the transmitter and receiver, together with the geometric configuration of the receiver's antenna [92, 172, 178]. The transmitter Tx is located at the origin while the receiver Rx is located at (x_R, y_R, z_R) or, equivalently, at (R_{RT}, θ, ϕ) in spherical coordinates. The receiver and the transmitter velocities are assumed to be horizontal. The velocity vector of the transmitter $\vec{v}_T$ is aligned with the x-axis

$$\vec{v}_T = v_T(1, 0, 0). \tag{3.1}$$

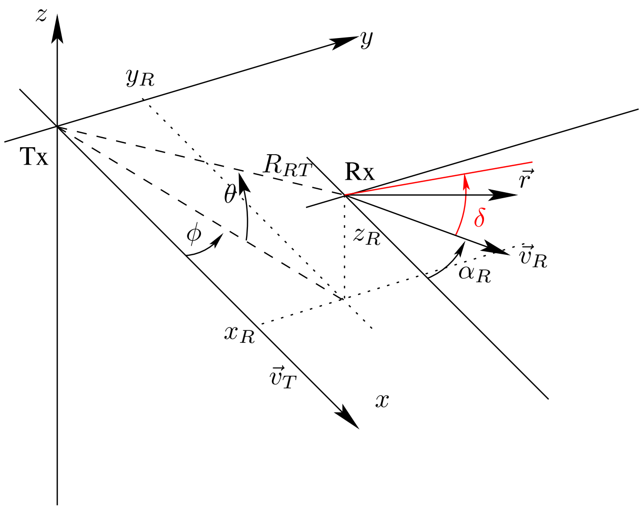


Figure 3.1: Bistatic transmitter/receiver geometric configuration.

The velocity vector of the receiver $\vec{v}_R$ makes an angle α_R w.r.t. the x-axis

$$\vec{v}_R = v_R(\cos\alpha_R, \sin\alpha_R, 0). \tag{3.2}$$

The reference axis of the receiving array lays in the xy-plane and makes an angle δ with $\vec{v}_R$. The angle δ is sometimes referred to as the "crab" angle in the case of a uniform linear array (ULA). The vector $\vec{r}$ describes the location of the considered antenna element with respect to an arbitrary reference point, for instance the mechanical center of the antenna array.

The ground is assumed flat and located at $-H$ with $H > 0$.

3.2.2 Bistatic scatterer geometric configuration

Figure 3.2 illustrates the scatterer geometry w.r.t. the transmitter and the receiver. The location of the scatterer S w.r.t. the transmitter Tx and the receiver Rx is characterized by the angles θ_T and ϕ_T and the angles θ_R and ϕ_R, respectively. The distances between the scatterer and both the receiver and the transmitter are denoted by R_R and R_T respectively. The location of the scatterer w.r.t. the transmitter is given by

$$R_T\vec{1}_{TS} \tag{3.3}$$

with $\vec{1}_{TS}$ a unit vector from Tx to S

$$\vec{1}_{TS} = (\cos\theta_T\cos\phi_T, \cos\theta_T\sin\phi_T, -\sin\theta_T). \tag{3.4}$$

Similarly, the location of the scatterer w.r.t. the receiver is given by

$$R_R\vec{1}_{RS} \tag{3.5}$$

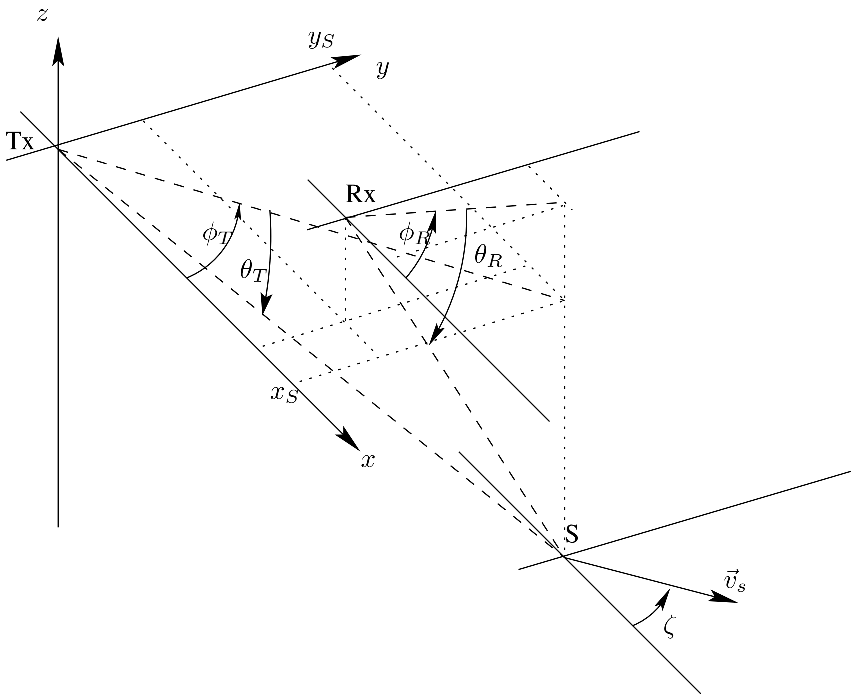


Figure 3.2: Bistatic scatterer geometric configuration

with $\vec{1}_{RS}$ a unit vector from Rx to S

$$\vec{1}_{RS} = (\cos\theta_R \cos\phi_R, \cos\theta_R \sin\phi_R, -\sin\theta_R). \tag{3.6}$$

The scatterer velocity vector is assumed laying in the xy-plane and makes an angle ζ with the x-axis

$$\vec{v}_S = v_S(\cos\zeta, \sin\zeta, 0). \tag{3.7}$$

3.3 Key bistatic parameters

3.3.1 Spatial frequency

This section establishes the expression of the spatial frequency and also deduces the expression for the electric field at the antenna elements.

The phase shift between the transmitted and the received signal at the antenna element located at $\vec{r}$ is given by $\phi(\vec{r}) = \frac{\tau_{rt}+\tau'(\vec{r})}{T_c} = \phi(0) + \frac{\tau'(\vec{r})}{T_c}$ where $\tau_{rt} = R/c$ is the round-trip time from the transmitter to the receiver antenna's center. The carrier frequency is denoted by f_c and $T_c = 1/f_c$ is the period of the carrier. $\phi(0) = \tau_{rt}/T_c$ is the phase reference corresponding to the center of the receiving antenna array. $\tau'(\vec{r})$ is the additional delay corresponding to the additional distance that the signal has to travel to reach the element located at $\vec{r}$. This delay is equal to [173]

$$\tau'(\vec{r}) = -\frac{\vec{r} \cdot \vec{1}_{RS}}{c}. \tag{3.8}$$

By considering only the factor depending on $\vec{r}$, the electric field at $\vec{r}$ is proportional to $e^{-j2\pi \frac{\tau'(\vec{r})}{T_c}}$, which can be rewritten as $E(\vec{r}) \propto e^{-j\vec{k}\cdot\vec{r}}$ thus defining the spatial wave vector $\vec{k}$ as [171]

$$\vec{k} = -2\pi \vec{1}_{RS} \frac{f_c}{c} = -\vec{1}_{RS} \frac{2\pi}{\lambda_c}. \tag{3.9}$$

This wave vector depends on the direction of arrival $\vec{1}_{RS}$ of the incident wave and on the wavelength $\lambda_c = \frac{c}{f_c}$. Thus, a spatial variation of the received signal (phase) appears at different positions on the receiving antenna due to the small time-delay evolution existing along the antenna.

Defining the normalized spatial frequency vector[1] as

$$\vec{\nu}_s = \vec{1}_{RS}/2 \tag{3.10}$$

yields $\vec{k} = -2\pi \vec{\nu}_s \frac{2}{\lambda_c}$ and normalizing the antenna element position in terms of half wavelength yields $\vec{\rho} = \frac{2}{\lambda_c}\vec{r}$. With these notations, the electric field in $\vec{r}$ can be rewritten as $E(\vec{\rho}) \propto e^{j2\pi\vec{\nu}_s\cdot\vec{\rho}}$.

3.3.2 Doppler frequency

Let us consider a transmitter transmitting a wave illuminating a scatterer S. The electric field at the receiving antenna due to the echo reflected by a scatterer S is proportional to $e^{j2\pi(f_c+f_D)t}$ where f_D is the Doppler frequency due to the transmitter, receiver and scatterer movements. After demodulation, the received signal is proportional to $e^{j2\pi f_D t}$. In bistatic scenarios, the Doppler frequency f_D is the sum of two terms: $f_D = f_{D_T} + f_{D_R}$, where f_{D_T} is the Doppler frequency due to the relative radial velocity of the transmitter w.r.t. the scatterer and f_{D_R} is the Doppler frequency due to the relative radial velocity of the receiver w.r.t. the scatterer. These two terms have the expressions

$$\begin{aligned} f_{D_T} &= \frac{1}{\lambda_c}(\vec{v}_T - \vec{v}_S)\vec{1}_{TS} \\ f_{D_R} &= -\frac{1}{\lambda_c}(\vec{v}_S - \vec{v}_R)\vec{1}_{RS} \end{aligned} \tag{3.11}$$

hence

$$\begin{aligned} f_{D_T} &= \frac{1}{\lambda_c}(v_T \cos\theta_T \cos\phi_T - v_S \cos\theta_T \cos(\zeta - \phi_T)) \\ f_{D_R} &= \frac{1}{\lambda_c}(v_R \cos\theta_R \cos(\alpha_R - \phi_R) - v_S \cos\theta_R \cos(\zeta - \phi_R)). \end{aligned} \tag{3.12}$$

[1]The normalized spatial frequency vector will be analyzed in detail in Chapter 5.

If the scatterer is stationary, $v_S = 0$ and one has

$$
\begin{aligned}
f_{D_T} &= \frac{1}{\lambda_c} v_T \cos\theta_T \cos\phi_T \\
f_{D_R} &= \frac{1}{\lambda_c} v_R \cos\theta_R \cos(\alpha_R - \phi_R).
\end{aligned} \tag{3.13}
$$

Assuming that the demodulated received signal is sampled at a frequency f_p, and normalizing the Doppler frequency w.r.t. this frequency yields

$$
\nu_D = \frac{f_{D_T} + f_{D_R}}{f_p} \tag{3.14}
$$

where ν_D is the normalized Doppler frequency expressed as a fraction of the sampling frequency. The m-th sample of the received signal is thus proportional to $e^{j2\pi m\nu_D}$.

3.3.3 Range

The bistatic range R is the total distance traveled by the signal from the transmitter to the scatterer to the receiver's antenna reference point. It is directly linked to the round-trip time $\tau_{rt} = R/c$. The range-rate is the derivative of this distance with respect to the time.

Isorange surfaces are surfaces where the bistatic range is constant, or, equivalently, surfaces of equal round-trip time. Isoranges are an important concept in radar as the signal corresponding to one particular range is the resultant of the contribution of all scatterers located on the isorange surface associated with the range of interest. In particular, ground clutter is the echo signal from distributed scatterers located on the ground. The ground clutter patches contributing to the signal at some specific range of interest will be located along an isorange curve which is the intersection of the isorange surface with the ground surface.

The equation for the isorange curve in the case of a flat ground surface is derived in Appendix A. In this case, the isoranges are ellipses.

Obviously, scatterers traveling with a velocity tangential to the isorange will exhibit a zero range-rate [174].

3.4 Space-time data model

For the sake of clarity of the text, let us consider a single transmitter transmitting a coherent pulse train. Chapter 6 extends the model to arbitrary signals. Let us consider P receive platforms. Platform p carries N_p receive channels (antenna elements) with arbitrary but known locations $\vec{\rho}_{p,n}$, $0 \leq n < N_p$. There is thus a total of $N = \sum_{p=1}^{P} N_p$ receive channels. At each receive channel, the signal is sampled in order to produce M samples in time corresponding to the M pulses transmitted during coherent processing interval (CPI). The sampling on each platform is synchronized such that the responses of the target arrive in phase at a reference channel on each platform, to compensates the decorrelation due to the time delay between the arrivals of the signals at all platforms [60].

3.4.1 Bistatic data snapshot

First, let us consider a single receive platform. For a platform p and for each of the M pulses during one coherent processing interval, one sample is obtained at each of the N_p elements of the antenna array. The bistatic range associated with the measurement is proportional to the time delay between the emission of the signal at the transmitter and the subsequent reception of the signal at the receiver.

Thus, M pulses results in N_pM space-time measurements at each platform p and at each range. These N_pM samples are called a snapshot. The notation $y_p(n, m)$, where $0 \leq n < N_p$ and $0 \leq m < M$, denotes the sample obtained at the n-th antenna element of platform p and corresponding to the m-th pulse. This operation can be repeated at each range, leading to a 3D datacube as illustrated in Figure 3.3 [173].

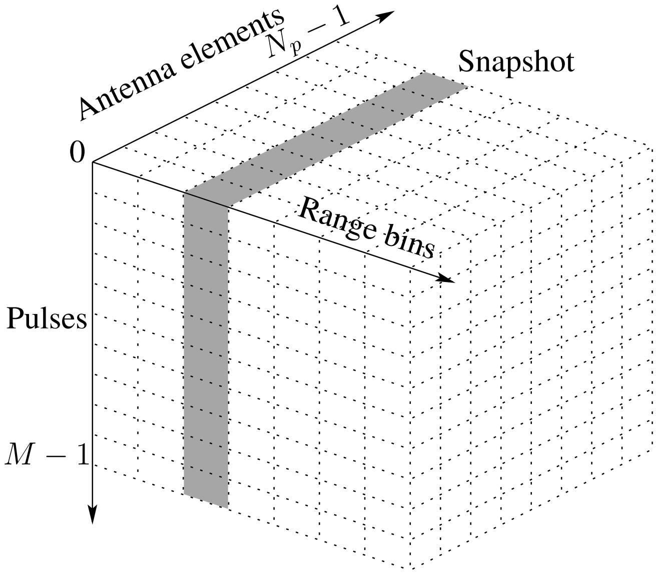


Figure 3.3: Illustration of the space-time datacube and a space-time snapshot.

The samples of the snapshot y_p can be lexicographically ordered yielding the vector

$$\mathbf{y}_p = [y_p(0,0), y_p(1,0), \ldots, y_p(N_p-1,0), y_p(0,1), \ldots, y_p(N_p-1,1), \\ \ldots, y_p(0, M-1), \ldots, y_p(N_p-1, M-1)]^T. \quad (3.15)$$

3.4.2 Bistatic space-time steering vector

According to Sections 3.3.1 and 3.3.2, the snapshot sample $y_p(n, m)$, at antenna element n and corresponding to pulse m, due to an isolated scatterer can be expressed as

$$y_p(n, m) \propto e^{j2\pi(\vec{\rho}_{p,n} \cdot \vec{\nu}_{p_s} + m\nu_{p_D})} \quad (3.16)$$

where $\propto$ denotes a proportionality relation. This only holds in the case of identical antenna element radiation patterns at the receiver, a uniform temporal sampling, and constant velocities of the transmitter, receiver, and scatterer during the coherent processing interval.

The corresponding vector $\mathbf{y}_p$ is thus proportional to a steering vector $\mathbf{v}$ which is itself the Kronecker product of a temporal steering vector $\mathbf{b}$ and a spatial steering vector $\mathbf{a}$ [60, 173]

$$\mathbf{v}(\vec{\nu}_{p_s}, \nu_{p_D}) = \mathbf{b}(\nu_{p_D}) \otimes \mathbf{a}(\vec{\nu}_{p_s}) \tag{3.17}$$

where

$$\mathbf{a} = [e^{j2\pi\vec{\rho}_{p,0}\cdot\vec{\nu}_{ps}} \; e^{j2\pi\vec{\rho}_{p,1}\cdot\vec{\nu}_{ps}} \; \cdots \; e^{j2\pi\vec{\rho}_{p,N_p-1}\cdot\vec{\nu}_{ps}}]^T \tag{3.18}$$

is the spatial steering vector,

$$\mathbf{b} = [1 \; e^{j2\pi\nu_{pD}} \; e^{j2\pi 2\nu_{pD}} \; \cdots \; e^{j2\pi(M-1)\nu_{pD}}]^T \tag{3.19}$$

is the temporal steering vector and $\otimes$ denotes the Kronecker product [34, Chapter 2] of its two vector arguments, $\vec{\nu}_{p_s}$ and ν_{p_D} are respectively the normalized spatial and Doppler frequencies associated with the isolated scatterer.

The vector $\mathbf{v}(\vec{\nu}_{p_s}, \nu_{p_D})$ is called the spatio-temporal steering vector or simply the steering vector associated with the isolated scatterer. To shorten the notation and to acknowledge the dependence of the steering vector on the considered platform, let us use the following equivalent notations for the steering vector associated with platform p

$$\mathbf{v}_p = \mathbf{v}_p(\boldsymbol{\theta}_p(\vec{x})) = \mathbf{v}_p(\vec{\nu}_{p_s}, \nu_{p_D}) \tag{3.20}$$

where the parameter vector $\boldsymbol{\theta}_p(\vec{x})$ contains the position and velocity vector of the transmit and receive platform, the location ($\vec{x}$) and velocity vector of the scatterer, the receive element positions $\vec{\rho}_{p,n}$, $0 \leq n < N_p$ and the parameters of the temporal sampling. The notation $\boldsymbol{\theta}_p(\vec{x})$ denote the vector of parameters for a steering vector corresponding to a scatterer located at $\vec{x}$. Finally, note that in a multistatic setup, the vector of parameters $\boldsymbol{\theta}_p$ is typically different for each bistatic radar, this dependence being indicated by the subscript p.

3.4.3 Multistatic space-time snapshot and steering vector

The multistatic space-time snapshot is obtained by stacking the bistatic snapshot $\mathbf{y}_p$ obtained from each platform and is noted [122]

$$\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T, \ldots, \mathbf{y}_P^T]^T. \tag{3.21}$$

3.5 Signal model

3.5.1 Target model

In our simplified model, targets are nothing else than point scatterers in the sense that we assume that the target extent is smaller than the resolution of the radar in range, azimuth angle, and Doppler.

3.5.1.1 Target reflectivity

Bistatic targets radar cross section (RCS) models are hardly discussed in the open literature. For instance the discussions in [174, 175] concentrate on the relationship of the monostatic RCS value with respect to the bistatic RCS for different bistatic angles and different frequencies. We are however not directly interested in the value of the RCS for a particular target, but rather in its temporal distribution, i.e., a modeling of the statistical fluctuation of the RCS. Let us discuss briefly the existing monostatic fluctuating target RCS models and assume that these extend to the bistatic case. There exists several models for the target complex random amplitude [120, 158], denoted α. Two of them are presented hereafter.

The Marcum model considers non fluctuating targets for which the value of α is constant. In monostatic configurations, this model is exact for a sphere. A variant is used in [22] where a target is modeled with a constant RCS and a phase ϕ uniformly distributed between 0 and 2π.

The Swerling-I and -II models assume the probability density function (PDF) of the target squared magnitude $|\alpha|^2$ is a negative exponential [120]

$$p(\alpha) = \frac{1}{\pi\sigma_\alpha^2} e^{-\frac{|\alpha|^2}{\sigma_\alpha^2}} \tag{3.22}$$

which results from the assumption that α is complex circular Gaussian distributed, i.e., $\alpha = \alpha_r + j\alpha_i$ where α_r and α_i are independent and normally distributed with zero mean and variance equal to $\frac{\sigma_\alpha^2}{2}$. In this case, the magnitude is Rayleigh distributed [132].

These models are applicable when the target can be modeled as a random assembly of independent scatterers, without any dominant one [120].

The Swerling-I model assumes slow fluctuations, i.e., the target complex amplitude is considered constant during the coherent processing interval (CPI). On the contrary, the Swerling-II model assumes random fluctuations during the CPI. These two cases can be seen as the two extremes, as during the CPI, the Swerling-I assumes full correlation while the Swerling-II model considers no correlation at all. Intermediate cases with temporal correlation of the target complex amplitudes are considered in [39].

Although other models exist [120], we will focus on two latter model families for the two following reasons: they are practically relevant — at least in the monostatic case — while providing results that remain analytically tractable.

Another issue in target modeling is the evolution of the RCS as a function of the look angle. The monostatic RCS usually varies a lot and exhibits deep fades as the look angle changes [158]. These variations are usually modeled by considering statistically independent complex amplitudes α at different look angles with different variances, which is referred to as *angular diversity*. On the contrary, a spatially coherent target is defined as a target with a completely correlated reflected complex amplitude at different look angles. The assumption of having coherent echoes from different look angles, although considered in the literature [76, 83] and possibly valid for particular targets, for instance targets consisting of a single dominant scatterer, is not realistic for real targets [14].

3.5.1.2 Bistatic target model

Under the assumption of a constant target velocity during the CPI, the contribution of one target to a bistatic snapshot is modeled as

$$\mathbf{y}_{t_p} = \alpha_p c_p \mathbf{v}_p(\vec{\nu}_{p_s}, \nu_{p_D}) \tag{3.23}$$

where $\mathbf{v}_p$ is the target's steering vector for radar p, α_p is the complex random amplitude of the target and c_p takes into account the factors of the radar equation and in particular the emission and reception antenna radiation patterns. $\vec{\nu}_{p_s}$ depends on the actual location of the target and is given by (3.10) while ν_{p_D} depends both on the location and on the velocity of the target and is given by (3.14).

Introducing the non-normalized steering vector

$$\mathbf{s}_p = c_p \mathbf{v}_p, \tag{3.24}$$

which takes into account the radiometric factors c_p of the range equation, we have

$$\mathbf{y}_{t_p} = \alpha_p \mathbf{s}_p. \tag{3.25}$$

3.5.1.3 Multistatic target model

The contribution of a point target to a multistatic snapshot results from the combination of the contribution of the target to several bistatic snapshots [122], i.e.,

$$\mathbf{y}_t = [\mathbf{y}_{t_1}^T, \mathbf{y}_{t_2}^T, \dots, \mathbf{y}_{t_P}^T]^T. \tag{3.26}$$

If a coherent target model is assumed, i.e., if the target is coherent among the different bistatic radars, one can rewrite (3.26) as

$$\mathbf{y}_{t_{\text{coherent}}} = \alpha[c_1\mathbf{v}_1(\boldsymbol{\theta}_1)^T, c_2\mathbf{v}_2(\boldsymbol{\theta}_2)^T, \dots, c_P\mathbf{v}_P(\boldsymbol{\theta}_P)^T]^T. \tag{3.27}$$

where $\alpha = \alpha_1 = \dots = \alpha_P$ is a random quantity behaving according to one of the models described in Section 3.5.1.1.

A more realistic target model implies that the complex target amplitudes α_p, as measured by different bistatic radars, are independent and possibly with different variance $E\{|\alpha|^2\}$ [14]. Under this condition, the contribution of a target to a multistatic snapshot (3.26) must be rewritten as

$$\mathbf{y}_{t_{\text{incoherent}}} = [\alpha_1 c_1\mathbf{v}_1(\boldsymbol{\theta}_1)^T, \alpha_2 c_2\mathbf{v}_2(\boldsymbol{\theta}_2)^T, \dots, \alpha_P c_P\mathbf{v}_P(\boldsymbol{\theta}_P)^T]^T. \tag{3.28}$$

where the α_p are statistically independent with variance $\sigma_{\alpha_p}^2$.

Introducing the non-normalized steering vector $\mathbf{s}_p$ yields

$$\mathbf{y}_{t_{\text{incoherent}}} = [\alpha_1\mathbf{s}_1^T, \alpha_2\mathbf{s}_2^T, \dots, \alpha_P\mathbf{s}_P^T]^T \tag{3.29}$$

and

$$\mathbf{y}_{t_{\text{coherent}}} = \alpha\mathbf{s} \tag{3.30}$$

with $\mathbf{s} = [\mathbf{s}_1^T, \mathbf{s}_2^T, \dots, \mathbf{s}_P^T]^T$.

3.5.2 Clutter model

3.5.2.1 Bistatic clutter model

First, let us consider a particular bistatic radar, say radar p. The return at a particular range is the integral of the signal reflected by the ground located along the corresponding isorange. Thus the contribution of clutter to the snapshot can be expressed as

$$\mathbf{y}_{c_p} = \int_A \alpha_p(\vec{x}) c_p(\vec{x}) \mathbf{v}_p(\vec{\nu}_{p_s}(\vec{x}), \nu_{p_D}(\vec{x})) d\vec{x} \tag{3.31}$$

where the integral is performed over the annulus A where there is a significant return. The width of this annulus depends on the range ambiguity function of the transmitted signal, as discussed in Chapter 6. The coefficient $c_p(\vec{x})$ represents the "amplitude" factors of the radar equation, including specifically the bistatic range factors together with the transmit and the receive antenna radiation patterns which depend on the look angle of the considered point $\vec{x}$ and the locations of the transmitter and the receiver.

This integral can be approached by a Riemann sum [60, 173] for which the annulus is divided in a large number of independent clutter patches evenly distributed in azimuth along the isorange

$$\mathbf{y}_{c_p} = \sum_{k=1}^{K} \alpha_{p,k} c_{p,k} \mathbf{v}_p(\vec{\nu}_{p_s,k}, \nu_{p_D,k}) \tag{3.32}$$

where $c_{p,k}$ contains all the geometric coefficients of the radar equation including the clutter patch area but not the RCS of that clutter patch; $\vec{\nu}_{p_s,k}$ and $\nu_{p_D,k}$ are respectively the reduced spatial and temporal frequencies at the center of clutter patch k and $\mathbf{v}_p(\vec{\nu}_{p_s,k}, \nu_{p_D,k})$ is the steering vector of the k-th clutter patch. $\alpha_{p,k}$ is the random complex amplitude of the scattered wave coming from clutter patch k. The number K of clutter patches to consider, briefly discussed in [21], must be at least larger than the dimension of the steering vectors $\mathbf{v}_p$, i.e., $K > N_p M$. Radar echoes typically exhibit a large dynamics, which may imply the use of a large number of clutter patches in order to correctly represent the clutter signal [149].

3.5.2.2 Clutter radar cross section

Under realistic conditions, there is a large number of individual scatterers contributing to the return signal coming from a single clutter patch [132]. Hence, the Swerling-I model given by (3.22) and discussed in a previous section can be assumed to correctly model the statistical behavior of the complex coefficient $\alpha_{p,k}$. Although the validity of this model over grass land is generally accepted, the assumptions of the scattering due to many individual scatterers is doubtful in urban environments. Models applicable in an urban context are discussed in [132]. The coefficients $\alpha_{p,k}$ are regarded as independent from a clutter patch to another, as the scatterers contributing to different $\alpha_{p,k}$ are different.

The power of the return from clutter patch k alone is given by $\sigma^2_{p_k} = E\{|\alpha_{p,k}|^2\}$ and is proportional to its area $A_{p,k}$. Dividing this returned power by the area $A_{p,k}$ yields the normalized scattering coefficient $\sigma_{0_{p_k}} = \sigma^2_{p_k}/A_{p,k}$ also called sigma naught which depends on the incidence angle. In monostatic cases, the constant-γ model

$$\sigma_0 = \gamma_0 \sin\theta, \tag{3.33}$$

where θ is the monostatic incidence angle and γ_0 a normalized reflectivity parameter, can model appropriately certain ground covers, e.g., the tropical rain forest [96]. In the latter case, an isotropic ground cover is assumed, as the look angle ϕ does not appear in expression (3.33). Over sea [68, 69] and over ice [27, 35], other types of models must be considered. Most of these models are empirical and are only valid in very limited cases, e.g., for particular ranges of incidence angles, at particular frequencies and polarizations. The monostatic constant-γ model, was adapted to the bistatic case in [174]

$$\sigma_0 = \gamma_0 \sin[(\theta_T + \theta_R)/2] \tag{3.34}$$

where θ_T is the incidence angle at the transmitter and θ_R the incidence angle at the receiver as shown on Figure 3.4. Unfortunately, these simple analytical models typically fail to render

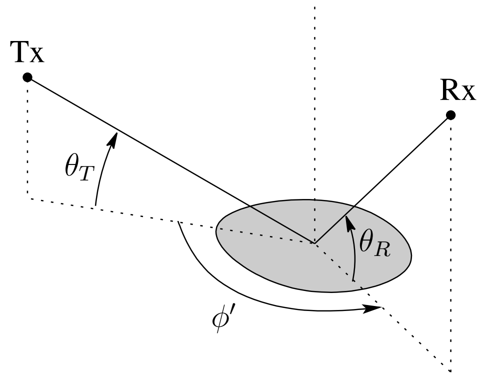


Figure 3.4: Bistatic measurement geometry.

the complexity of real world situations, as shown in the bistatic clutter databases presented in [174, 175]. A general trend is however apparent in [174]: there is a minimum of σ_0 at an out-of-incidence plane angle $\phi' = 90^o$ and a maximum at an in-incidence plane angle $\phi' = 180^o$ (specular reflexion). Similarly, in elevation, a maximum corresponding to specular reflexion ($\theta_T = \theta_R$) is obtained.

3.5.2.3 Multistatic clutter model

As for the multistatic target model, the clutter contribution to a multistatic snapshot results from the combination of the bistatic clutter contributions of the different bistatic radars, i.e.,

$$\mathbf{y}_c = [\mathbf{y}^T_{c_1}, \mathbf{y}^T_{c_1}, \ldots, \mathbf{y}^T_{c_P}]^T \tag{3.35}$$

where $\mathbf{y}_{c_p}$ denotes the bistatic clutter contribution of the bistatic radar p.

The different bistatic radars have generally different spatial locations, i.e., the receivers are not collocated. Hence, the isoranges of different bistatic radars are generally not overlapping and the clutter signal component present in the space-time samples of different receive platforms are independent[2] as noted in [24].

3.5.2.4 Multistatic clutter covariance matrix

As the complex clutter reflectivity is assumed zero mean circularly Gaussian distributed, the clutter snapshot $\mathbf{y}_c$ is a zero mean multidimensional complex Gaussian random vector, and its distribution is completely characterized by its covariance matrix.

The multistatic clutter covariance matrix is defined as

$$\mathbf{R}_c = E\{\mathbf{y}_c\mathbf{y}_c^\dagger\}. \tag{3.36}$$

From this definition, one immediately deduces that the covariance matrix is Hermitian $\mathbf{R}_c = \mathbf{R}_c^\dagger$. The PDF of the snapshot is given by

$$p(\mathbf{y}_c) = \frac{1}{\pi^{NM}|\mathbf{R}_c|} e^{-\mathbf{y}_c^\dagger \mathbf{R}_c^{-1}\mathbf{y}_c} \tag{3.37}$$

where $|\mathbf{R}_c|$ denotes the determinant of the matrix $\mathbf{R}_c$. In the same way, as the clutter snapshot signal from each bistatic radar is assumed zero mean circularly Gaussian distributed, it follows that

$$p(\mathbf{y}_{c_p}) = \frac{1}{\pi^{N_pM}|\mathbf{R}_{c_p}|} e^{-\mathbf{y}_{c_p}^\dagger \mathbf{R}_{c_p}^{-1}\mathbf{y}_{c_p}} \tag{3.38}$$

where

$$\mathbf{R}_{c_p} = E\{\mathbf{y}_{c_p}\mathbf{y}_{c_p}^\dagger\}. \tag{3.39}$$

is the clutter covariance matrix of the clutter seen by bistatic radar p. The independence of the $\mathbf{y}_{c_p}$'s yields

$$p(\mathbf{y}_c) = \prod_{p=1}^{P} p(\mathbf{y}_{c_p}) \tag{3.40}$$

and hence

$$p(\mathbf{y}_c) = \prod_{p=1}^{P} \frac{1}{\pi^{N_pM}|\mathbf{R}_{c_p}|} e^{-\mathbf{y}_{c_p}^\dagger \mathbf{R}_{c_p}^{-1}\mathbf{y}_{c_p}}. \tag{3.41}$$

Independence of the clutter snapshots from different bistatic radars also imply that the covariance matrix $\mathbf{R}_c$ has a block-diagonal structure

$$\mathbf{R}_c = \begin{bmatrix} \mathbf{R}_{c_1} & & & 0 \\ & \mathbf{R}_{c_2} & & \\ & & \ddots & \\ 0 & & & \mathbf{R}_{c_P} \end{bmatrix} \tag{3.42}$$

[2]Possible dependence where the isoranges intersect is neglected as the look-angle is in general very different, which causes decorrelation.

where the p^{th} block on the diagonal corresponds to the clutter covariance matrix $\mathbf{R}_{c_p}$ of the clutter seen by the p^{th} bistatic radar. Taking the particular structure of the covariance matrix into account and noting that the determinant of a block diagonal matrix is the product of the determinant of the matrices on the diagonal

$$|\mathbf{R}_c| = \prod_{p=1}^{P} |\mathbf{R}_{c_p}|. \tag{3.43}$$

and that the inverse of a block diagonal matrix is a block diagonal matrix composed of the inverse of each blocks on its diagonal, the equality between (3.37) and (3.41) follows.

3.5.3 Overall signal model

Besides target and clutter returns, the space-time snapshot also includes some thermal noise $\mathbf{y}_n$, assumed white circularly Gaussian, uncorrelated with the other signal components, and with covariance matrix $\sigma_n^2\mathbf{I}$ where $\mathbf{I}$ is the $NM \times NM$ identity matrix.

Hence the overall signal model is given by

$$\mathbf{y} = \mathbf{y}_t + \mathbf{y}_{i+n} \tag{3.44}$$

where

$$\mathbf{y}_{i+n} = \mathbf{y}_c + \mathbf{y}_n \tag{3.45}$$

will be referred to as the interference-plus-noise signal and has a covariance matrix

$$\mathbf{R}_{i+n} = \mathbf{R}_c + \sigma_n^2\mathbf{I}. \tag{3.46}$$

In the rest of this document and when no confusion is possible, the subscript $i + n$ is dropped and the interference plus noise covariance matrix is simply denoted by $\mathbf{R}$ (and $\mathbf{R}_p$ if the bistatic radar p is considered).

3.6 Summary

Starting from the acquisition geometry, the expressions for the Doppler frequency and the spatial frequency were introduced for a bistatic radar to define the space-time steering vector used to model a point target. Models able to render the spatial (angular) correlation of the target returns were presented and serves as the basis for the discussion of the detection theory in the next chapter. Finally, the clutter model was introduced together with the interference-plus-noise covariance matrix in the multistatic case.

Chapter 4

Multistatic target detection

4.1 Introduction

This chapter reviews the detection theory and applies it to multistatic radars in order to study their detection performance and in particular to investigate the conditions under which a multistatic configuration is more advantageous than a bistatic one.

The considered multistatic radar consists of several bistatic radars, the performance analysis of bistatic radars serves as basis for the study of the performance of multistatic radars and is performed with the Marcum and Swerling-I target models described in the previous chapter.

The detection is shown to be equivalent to thresholding the squared magnitude of the output of a filter applied to the received signal. The detection performance is shown to be monotonically increasing with the signal-to-interference-plus-noise ratio (SINR). This result is then used to assess the performance of different filters.

4.2 Detection theory

Target detection can be seen as a binary composite hypothesis testing, as the PDF in the presence of a target (hypothesis H_1) depends on some unknown parameters. Under hypothesis H_0, no target is present and the received signal can be modeled as

$$\mathbf{y}_p = \mathbf{y}_{i+n_p} \tag{4.1}$$

where $\mathbf{y}_{i+n_p}$ is the interference-plus-noise signal as modeled in Section 3.5.2. Let us assume complete knowledge of the interference-plus-noise covariance matrix $\mathbf{R}_p$, i.e., the clairvoyant case, in order to develop the detectors and study their performance. The case of the unknown interference-plus-noise covariance matrix is discussed in Section 4.5. Under hypothesis H_1, a target is present with (assumed) known steering vector $\mathbf{s}_p$, and the received signal can be modeled as

$$\mathbf{y}_p = \alpha_p \mathbf{s}_p + \mathbf{y}_{i+n_p} \tag{4.2}$$

where α_p is the complex amplitude[1] of the target return. Let us also assume complete knowledge of the target's steering vector. The latter hypothesis is discussed later in Section 4.5.

Two main criteria are used in binary hypothesis testing [82, 155, 169]: the Bayes criterion and the Neyman-Pearson criterion. The Bayes criterion requires knowledge of *a priori* probabilities for each hypothesis, the definition of a cost associated with each possible decision and minimizes that cost. The Neyman-Pearson criterion bypasses the difficulty of having to define realistic costs and *a priori* probabilities that can arise when dealing with physical situations and is based on the maximization of the probability of detection for a given probability of false alarm. Both criteria yield the same decision rule [82, 169].

4.3 Detection criterion

4.3.1 The likelihood ratio test

Let us consider a bistatic radar p. The PDF $p(\mathbf{y}_p; H_0)$ and $p(\mathbf{y}_p; H_1)$ are the probability densities of the measurements $\mathbf{y}_p$ under the two considered hypotheses, H_0 and H_1, respectively. If Gaussian probability density functions (PDF) are assumed, then [54, 81]

$$p(\mathbf{y}_p; H_0) = \frac{1}{\pi^{N_pM}|\mathbf{R}_p|} e^{-\mathbf{y}_p^\dagger \mathbf{R}_p^{-1} \mathbf{y}_p} \tag{4.3}$$

and

$$p(\mathbf{y}_p|\alpha_p; H_1) = \frac{1}{\pi^{N_pM}|\mathbf{R}_p|} e^{-(\mathbf{y}_p - \alpha_p \mathbf{s}_p)^\dagger \mathbf{R}_p^{-1} (\mathbf{y}_p - \alpha_p \mathbf{s}_p)} \tag{4.4}$$

where N_pM is the number of elements of the vector $\mathbf{y}_p$, $\mathbf{R}_p$ is the interference-plus-noise covariance matrix

$$\mathbf{R}_p = E\{\mathbf{y}_{i+n_p}\mathbf{y}_{i+n_p}^\dagger\} \tag{4.5}$$

and $|\mathbf{R}_p|$ denotes the determinant of this matrix. The function $L(H_0) = p(\mathbf{y}_p; H_0)$ viewed as a property of H_0 given the measurements $\mathbf{y}_p$ is the likelihood of H_0. Similarly, $L(H_1) = p(\mathbf{y}_p; H_1)$. If α_p is random with known PDF $p(\alpha_p)$, the likelihood of H_1, $L(H_1)$, is obtained by integrating over the possible values of α_p taking into account its PDF

$$L(H_1) = \int_D p(\mathbf{y}_p, \alpha_p; H_1) d\alpha_p = \int_D p(\mathbf{y}_p|\alpha_p; H_1) p(\alpha_p) d\alpha_p \tag{4.6}$$

where D is the integration domain, generally $\mathbb{C}$.

A Neyman-Pearson test is based on thresholding the likelihood ratio

$$\Lambda(\mathbf{y}_p) = \frac{L(H_1)}{L(H_0)} = \frac{\displaystyle\int_D p(\mathbf{y}_p|\alpha_p; H_1) p(\alpha_p) d\alpha_p}{p(\mathbf{y}_p; H_0)} \tag{4.7}$$

[1]When a Marcum model is considered, the amplitude is deterministic and unknown and when a Swerling-I target model is considered, the amplitude is random.

The same result is obtained by using a Bayesian approach [82, 169]. The corresponding test is the likelihood ratio test (LRT)

$$\Lambda(\mathbf{y}_p) \underset{H_0}{\overset{H_1}{\gtrless}} \eta_0 \tag{4.8}$$

where η_0 is a threshold, generally data-dependent and governed by the desired probability of false alarm. The result of the test does not change if each side of (4.8) is transformed by using a monotonously increasing function. As the likelihood ratio generally involves exponentials, the logarithm of the likelihood is often used and yields the log-likelihood ratio test

$$\lambda(\mathbf{y}_p) \underset{H_0}{\overset{H_1}{\gtrless}} \eta_0' \tag{4.9}$$

where $\lambda(\mathbf{y}_p) = \ln \Lambda(\mathbf{y}_p)$ and $\eta_0' = \ln \eta_0$.

4.3.2 The generalized likelihood ratio test

In some particular cases, the likelihood $L(H_1)$ cannot be computed, since it depends on deterministic but unknown parameters or on stochastic parameters with unknown PDF, or simply, a closed-form of the integral (4.6) cannot be obtained. In these cases, a possible procedure [169] consists in maximizing the likelihood over the unknown parameters and using the values of these maxima in the LRT. This procedure is called the generalized likelihood ratio test, written here for a deterministic unknown α_p

$$\Lambda_g(\mathbf{y}_p) = \frac{\max\limits_{\alpha_p} p(\mathbf{y}_p; \alpha_p, H_1)}{p(\mathbf{y}_p; H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_0. \tag{4.10}$$

Note that this is, by definition, equivalent to replacing the unknown parameters by their maximum likelihood value. Obviously, the generalized log-likelihood is defined in a similar way as in the previous sub-section, e.g., $\lambda_g(\mathbf{y}_p) = \ln \Lambda_g(\mathbf{y}_p)$.

Although the GLRT is not an optimum test, nevertheless, it usually provides acceptable results in the considered detection problem [169]. Moreover, other methods can be used such as the Wald test or the Rao test [82].

4.3.3 Bistatic scenarios

This section derives the sufficient statistic in the case of a single bistatic radar p.

4.3.3.1 Marcum target model

In the case of a Marcum target model, α_p is a deterministic unknown parameter and the GLRT is used. The maximum likelihood estimation of α_p is obtained by setting the derivative of the logarithm of (4.4) with respect to α_p equal to zero, which yields (the detailed derivation is presented in Appendix C.2)

$$\hat{\alpha}_p = \frac{\mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{y}_p}{\mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{s}_p}. \tag{4.11}$$

Replacing this value in (4.10) yields the following expression for the generalized log-likelihood ratio test,

$$T_{b,m}(\mathbf{y}_p) = \frac{|\mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{y}_p|^2}{\mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{s}_p} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{b,m} \tag{4.12}$$

where $T_{b,m}$ groups the terms of the logarithm of the likelihood ratio that depend on $\mathbf{y}_p$. The threshold $\eta_{b,m}$ includes the additive constants that do not depend on $\mathbf{y}_p$. The subscripts b and m refer to *bistatic* radar and *Marcum* target model respectively. Expression (4.12) corresponds to the well known result in detection theory [114, 150] and is a modified version of the well known criterion initially proposed in [22, 142] to have constant false alarm rate (CFAR) properties. The absolute value appearing in (4.12) expresses the independence of the detection problem on the phase of $\hat{\alpha}_p$.

4.3.3.2 Swerling-I target model

In the case of a Swerling-I target model, the complex target amplitude is assumed random with known PDF

$$p(\alpha_p) = \frac{1}{\pi \sigma_{\alpha_p}^2} e^{-\frac{|\alpha_p|^2}{\sigma_{\alpha_p}^2}} \tag{4.13}$$

which implies the knowledge of $\sigma_{\alpha_p}^2$, the variance of α_p. In this case,

$$L(H_1) = p(\mathbf{y}_p; H_1) = \int_D p(\mathbf{y}_p | \alpha_p; H_1) p(\alpha_p) d\alpha_p \tag{4.14}$$

where D is the whole complex plane $\mathbb{C}$, and this yields (Appendix C.3 provides with the detailed derivation)

$$L(H_1) = p(\mathbf{y}_p; H_1) = \frac{1}{\pi^{N_p M} |\mathbf{R}_p|} \frac{1}{\pi \sigma_{\alpha_p}^2} \frac{\pi}{\sqrt{\frac{1}{\sigma_{\alpha_p}^2} + \mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{s}_p}} e^{-\mathbf{y}_p^\dagger \mathbf{R}_p^{-1} \mathbf{y}_p} e^{\frac{|\mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{y}_p|^2}{\frac{1}{\sigma_{\alpha_p}^2} + \mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{s}_p}}. \tag{4.15}$$

Taking into account (4.3), the likelihood ratio (4.7) is given by

$$\Lambda(\mathbf{y}_p) = \frac{1}{\pi \sigma_{\alpha_p}^2} \frac{\pi}{\sqrt{\frac{1}{\sigma_{\alpha_p}^2} + \mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{s}_p}} e^{\frac{|\mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{y}_p|^2}{\frac{1}{\sigma_{\alpha_p}^2} + \mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{s}_p}} \tag{4.16}$$

or, after taking the logarithm of the latter expression,

$$\lambda(\mathbf{y}_p) = \frac{|\mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{y}_p|^2}{\frac{1}{\sigma_{\alpha_p}^2} + \mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{s}_p} + \kappa \tag{4.17}$$

where κ is a deterministic constant that does not depend on $\mathbf{y}_p$. The corresponding log-likelihood ratio test is then

$$T_{b,s}(\mathbf{y}_p) = \frac{|\mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{y}_p|^2}{\frac{1}{\sigma_{\alpha_p}^2} + \mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{s}_p} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{b,s} \tag{4.18}$$

where the subscripts b and s refer to *bistatic* radar and *Swerling-I* target model respectively. Equivalently, rewriting this expression as

$$T_{b,s}(\mathbf{y}_p) = \frac{\mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{s}_p}{\frac{1}{\sigma_{\alpha_p}^2} + \mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{s}_p} \frac{|\mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{y}_p|^2}{\mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{s}_p} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{b,s} \tag{4.19}$$

and including the first term in the constant $\eta'_{b,s}$ yields

$$T'_{b,s}(\mathbf{y}_p) = \frac{|\mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{y}_p|^2}{\mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{s}_p} \underset{H_0}{\overset{H_1}{\gtrless}} \eta'_{b,s} \tag{4.20}$$

which is exactly the same expression as in the case of a Marcum target model (4.12). The test itself will only differ by the threshold.

4.3.3.3 Discussion

Obviously, there is a close relationship between the result (4.12) obtained by considering a deterministic unknown α_p and (4.18) obtained by considering a random α_p with known Gaussian PDF. In the random case, α_p is assumed to have a complex circular Gaussian distribution of variance $\sigma_{\alpha_p}^2$, which can be considered as *a priori* knowledge about α_p. If no *a priori* knowledge was assumed, which is equivalent to letting $\sigma_{\alpha_p} \to \infty$, (4.18) reduces to (4.12).

These detectors rewritten

$$|\mathbf{w}_p^\dagger \mathbf{y}_p|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \eta_b \tag{4.21}$$

are expressed as a function of an optimal (or nearly optimal in the case of a deterministic α_p) weight vector $\mathbf{w}_p$ given by

$$\mathbf{w}_p = k \mathbf{R}_p^{-1} \mathbf{s}_p. \tag{4.22}$$

where k is a constant. This weight vector corresponds to the generalized matched filter [82], well known in detection theory [22, 142, 171]. The quantity $|\mathbf{w}_p^\dagger \mathbf{y}_p|$ is called the sufficient statistic [169].

4.3.4 Multistatic scenario

In a multistatic scenario, the model described in the previous chapter, in the presence of a target, yields the following snapshot expression

$$\mathbf{y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_P]^T = [\alpha_1 \mathbf{s}_1 + \mathbf{y}_{i+n_1}, \alpha_2 \mathbf{s}_2 + \mathbf{y}_{i+n_2}, \ldots \alpha_P \mathbf{s}_P + \mathbf{y}_{i+n_P}]^T \tag{4.23}$$

where $\alpha_1, \ldots, \alpha_P$ are the complex amplitudes of the target echo signals received by the respective bistatic radars $1, \ldots, P$.

If the clutter contributions to the received signals of each bistatic radar are independent, the PDF of $\mathbf{y}$ if the target is not present, i.e., under hypothesis H_0, is given by

$$p(\mathbf{y}; H_0) = \prod_{p=1}^{P} p(\mathbf{y}_p; H_0) = \prod_{p=1}^{P} \frac{1}{\pi^{N_p M} |\mathbf{R}_p|} e^{-\mathbf{y}_p^\dagger \mathbf{R}_p^{-1} \mathbf{y}_p}. \tag{4.24}$$

The actual distribution of $\mathbf{y}$ if a target is present depends on the considered target model and is further detailed in the following sections together with the corresponding test statistic.

4.3.4.1 Marcum target model

If the deterministic unknown target reflectivity α_p is different for each bistatic radar, and given the independence of the interference-plus-noise of the different radars,

$$p(\mathbf{y}; \alpha_1, \ldots, \alpha_P, H_1) = \prod_{p=1}^{P} p(\mathbf{y}_p; \alpha_p, H_1) = \prod_{p=1}^{P} \frac{1}{\pi^{N_p M} |\mathbf{R}_p|} e^{-(\mathbf{y}_p - \alpha_p \mathbf{s}_p)^\dagger \mathbf{R}_p^{-1} (\mathbf{y}_p - \alpha_p \mathbf{s}_p)}. \tag{4.25}$$

The generalized likelihood ratio is given by

$$\Lambda_g = \frac{\max\limits_{\alpha_1, \ldots, \alpha_P} p(\mathbf{y}; \alpha_1, \ldots, \alpha_P, H_1)}{p(\mathbf{y}; H_0)} = \prod_{p=1}^{P} \frac{\max\limits_{\alpha_p} p(\mathbf{y}_p; \alpha_p, H_1)}{p(\mathbf{y}_p; H_0)} \tag{4.26}$$

and, from the result of Section 4.3.3.1, the generalized log-likelihood ratio test yields

$$T_{m,m}(\mathbf{y}) = \sum_{p=1}^{P} \frac{|\mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{y}_p|^2}{\mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{s}_p} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{m,m} \tag{4.27}$$

where the subscripts m and m refer to *multistatic* radar and *Marcum* target model respectively. The test statistic for the multistatic radar in the case of a Marcum target model and with the assumption of uncorrelated clutter contributions of different bistatic radars is thus simply the sum of the test statistics for each bistatic radar composing the multistatic radar.

4.3.4.2 Swerling-I target model

In the case of a Swerling-I target model, the different target echo signal complex amplitudes α_p of the different bistatic radars are random and statistically independent and their PDF are given by

$$p(\alpha_p) = \frac{1}{\pi \sigma_{\alpha_p}^2} e^{-\frac{|\alpha_p|^2}{\sigma_{\alpha_p}^2}} \tag{4.28}$$

where the σ_{α_p}'s are assumed to be known. As both α_p and the interference-plus-noise contribution $\mathbf{y}_{i+n_p}$ are assumed independent for each bistatic radar,

$$L(H_1) = p(\mathbf{y}; H_1) = \prod_{p=1}^{P} p(\mathbf{y}_p; H_1) \tag{4.29}$$

where $p(\mathbf{y}_p|H_1)$ is given by (4.15). A similar calculation as in Section 4.3.3.2 provides with the following test

$$T_{m,s}(\mathbf{y}) = \sum_{p=1}^{P} \frac{|\mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{y}_p|^2}{\frac{1}{\sigma_{\alpha_p}^2} + \mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{s}_p} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{m,s} \tag{4.30}$$

which can be rewritten as

$$T_{m,s}(\mathbf{y}) = \sum_{p=1}^{P} \frac{\mathrm{SINR}_{p,s}}{1 + \mathrm{SINR}_{p,s}} \frac{|\mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{y}_p|^2}{\mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{s}_p} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{m,s} \tag{4.31}$$

where $\mathrm{SINR}_{p,s} = \sigma_{\alpha_p}^2 \mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{s}_p$ is called the signal to interference-plus-noise ratio and the subscript s denotes that it is defined for the case of a Swerling-I target model. A similar expression[2] is found in [24] and for continuous time signals in white noise in [32].

As can be seen, the test statistic of the multistatic radar is a weighted sum of the test statistic of each bistatic radar (4.20). The weight assigned to the contribution of a particular bistatic radar depends on the SINR corresponding to that particular radar. The contributions of radars with a low SINR have a smaller weight in the sum.

4.3.5 Discussion

The detection statistic for a multistatic radar (4.27) and (4.30) in the case of a Marcum target model and a Swerling-I target model respectively is the incoherent sum of the contributions of each bistatic radar constituting the multistatic radar [24], as the relative phase of the complex amplitude α_p of the different radars is unknown.

4.4 Performance

4.4.1 Introduction

The performance of the test is measured by the probability of detection for a given probability of false alarm. If the test statistic is noted T and the test is unilateral

$$T \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \tag{4.32}$$

[2] The $\mathrm{SINR}_{p,s}$ is multiplied by a factor 2, owing to a different definition of $\sigma_{\alpha_p}^2$ (no further reference concerning the expression is given in that paper).

the probability of false alarm is given by

$$P_{\text{FA}} = Q(\eta) = \int_{\eta}^{+\infty} p(T; H_0) dT \tag{4.33}$$

where $p(T|H_0)$ is the PDF of the test statistic T under condition H_0, and $Q(\eta)$ the complement of the cumulative density function (CDF) of T under condition H_0. As the complement of the CDF is always a monotonously decreasing function, its inverse Q^{-1} always exists. This inverse function can be used to compute the threshold η from the desired probability of false alarm

$$\eta = Q^{-1}(P_{\text{FA}}). \tag{4.34}$$

Similarly, the probability of detection is given by

$$P_{\text{D}} = \int_{\eta}^{+\infty} p(T; H_1) dT \tag{4.35}$$

where $p(T; H_1)$ is the PDF of T under hypothesis H_1.

4.4.2 Bistatic scenarios

Again, let us first consider the detection performance of a single bistatic radar and subsequently extend the results to multistatic radars. Similar-looking results can be found in the literature: the Swerling-I continuous-time case is handled in [169]; the discrete case is handled in [82, 154], but only either for the known-signal in colored noise or the Swerling-I target in white noise. The case of complex signals combined with colored noise is not addressed for the Marcum and Swerling-I target models. The derivations in this section are however inspired from [82].

4.4.2.1 Marcum target model

From (4.12), the test statistics is $T_{b,m}$, which can be expressed as $T_{b,m} = |z_p|^2$ with

$$z_p = \frac{1}{\gamma_p} \mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{y}_p \tag{4.36}$$

where $\gamma_p = \sqrt{\mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{s}_p}$. The rationale behind this formulation is that the computation of the distribution of z_p and the subsequent deduction of the distribution of $T_{b,m}$ is easier than a direct computation of the distribution of $T_{b,m}$. Indeed, as z_p is a linear combination of complex Gaussian random variables, z_p is also a complex Gaussian random variable. Under hypothesis H_0, z_p is zero-mean

$$E\{z_p; H_0\} = \frac{1}{\gamma_p} \mathbf{s}_p^\dagger \mathbf{R}_p^{-1} E\{\mathbf{y}_{i+n_p}\} = 0 \tag{4.37}$$

and its variance is equal to unity

$$\operatorname{var}\{z_p; H_0\} = E\{z_p z_p^\dagger; H_0\} = \frac{1}{\gamma_p^2} \mathbf{s}_p^\dagger \mathbf{R}_p^{-1} E\{\mathbf{y}_{i+n_p} \mathbf{y}_{i+n_p}^\dagger\} \mathbf{R}_p^{-1} \mathbf{s}_p = \frac{1}{\gamma_p^2} \mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{s}_p = 1 \tag{4.38}$$

as $E\{\mathbf{y}_{i+n_p}\mathbf{y}_{i+n_p}^\dagger\} = \mathbf{R}_p$. Finally,

$$H_0 : z_p \sim \mathcal{CN}(0,1) \tag{4.39}$$

which implies $H_0 : \mathrm{Re}(z_p) \sim \mathcal{N}(0, \frac{1}{2})$ and $H_0 : \mathrm{Im}(z_p) \sim \mathcal{N}(0, \frac{1}{2})$. Hence, since $|z_p|^2 = \mathrm{Re}(z_p)^2 + \mathrm{Im}(z_p)^2$, $2|z_p|^2$ has a chi-squared PDF[3] with 2 degrees of freedom

$$H_0 : |z_p|^2 \sim \frac{\chi_2^2}{2} \tag{4.40}$$

which is also known as a negative exponential PDF.

Similarly, under hypothesis H_1, the expectation of z_p is

$$E\{z_p; H_1\} = \frac{1}{\gamma_p}\mathbf{s}_p^\dagger\mathbf{R}_p^{-1}\mathbf{s}_p\alpha_p + \frac{1}{\gamma_p}\mathbf{s}_p^\dagger\mathbf{R}_p^{-1}E\{\mathbf{y}_{i+n_p}\} = \gamma_p\alpha_p \tag{4.41}$$

and its variance is again equal to unity

$$\mathrm{var}\{z_p; H_1\} = E\{|z_p - \gamma_p\alpha_p|^2; H_1\} = \frac{1}{\gamma_p^2}\mathbf{s}_p^\dagger\mathbf{R}_p^{-1}E\{\mathbf{y}_{i+n_p}\mathbf{y}_{i+n_p}^\dagger\}\mathbf{R}_p^{-1}\mathbf{s}_p = \frac{1}{\gamma_p^2}\mathbf{s}_p^\dagger\mathbf{R}_p^{-1}\mathbf{s}_p = 1 \tag{4.42}$$

hence

$$H_1 : z_p \sim \mathcal{CN}(\gamma_p\alpha_p, 1). \tag{4.43}$$

From the previous equation, it follows that $2|z_p|^2$ has a non-central chi-squared PDF[4] with 2 degrees of freedom and non-centrality parameter $2\gamma_p^2|\alpha_p|^2$

$$H_1 : |z_p|^2 \sim \frac{1}{2}\chi_2'^2(2\gamma_p^2|\alpha_p|^2). \tag{4.44}$$

Finally, as $T_{b,m} = |z_p|^2$,

$$T_{b,m}(\mathbf{y}_p) \sim \begin{cases} \frac{1}{2}\chi_2^2 & H_0 \\ \frac{1}{2}\chi_2'^2(2\gamma_p^2|\alpha_p|^2) & H_1. \end{cases} \tag{4.45}$$

Figure 4.1 (a) illustrates the distribution of the test statistic $T_{b,m}$ under both hypotheses.

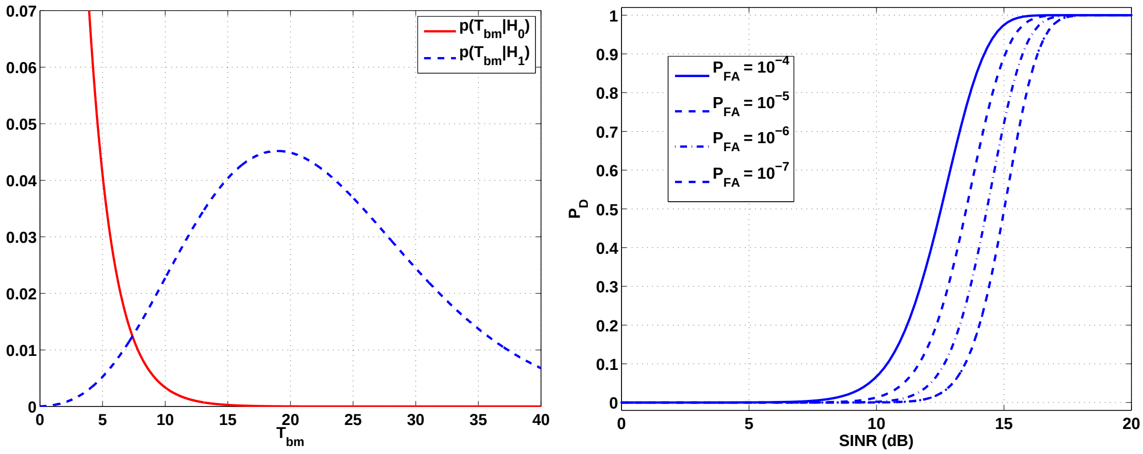
The probability of false alarm is obtained by computing

$$P_{\mathrm{FA}} = Q_{\chi_2^2}(2\eta_{b,m}) = e^{-\eta_{b,m}}. \tag{4.46}$$

where $Q_{\chi_\nu^2}$ is the complement of the CDF of the central chi-squared distribution with ν degrees of freedom. It is remarkable that the P_{FA} neither depends on the interference-plus-noise covariance matrix $\mathbf{R}_p$ nor on the steering vector normalization nor on the target reflectivity α_p, which

[3]The properties of this distribution are reviewed in Appendix B.

[4]The properties of this distribution are reviewed in Appendix B.



(a) (b)

Figure 4.1: (a) PDF of the test statistic under the two hypotheses and (b) probability of detection P_{D} as a function of the SINR for different probabilities of false alarm P_{FA}.

justifies the claim that the detector (4.12) has CFAR properties [150]. Moreover, defining the signal to interference-plus-noise ratio for the Marcum target model $\mathrm{SINR}_{p,m}$ as

$$\mathrm{SINR}_{p,m} = \gamma_p^2 |\alpha_p|^2 = \mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{s}_p |\alpha_p|^2 \tag{4.47}$$

and, using (4.46) yields

$$P_{\mathrm{D}} = Q_{\chi_2'^2(2\mathrm{SINR}_{p,m})}(2\eta_{b,m}) = Q_{\chi_2'^2(2\mathrm{SINR}_{p,m})}(-2\ln P_{\mathrm{FA}}) \tag{4.48}$$

where $Q_{\chi_\nu'^2(\mu)}$ is the complement of the CDF of the non-central chi-squared distribution with ν degrees of freedom and non-centrality parameter μ. As CDF $F_{\chi_2'^2(2\mathrm{SINR}_{p,m})} = 1 - Q_{\chi_2'^2(2\mathrm{SINR}_{p,m})}$ is decreasing with increasing non-centrality parameter $2\mathrm{SINR}_{p,m}$, the probability of detection increases with the SINR as is shown in Figure 4.1 (b).

4.4.2.2 Swerling-I target model

From (4.18), let us define

$$z_p = \frac{1}{\gamma_p} \mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{y}_p \tag{4.49}$$

where $\gamma_p = \sqrt{\frac{1}{\sigma_{\alpha_p}^2} + \mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{s}_p}$. As z_p is a linear combination of complex Gaussian random variables, z_p is also a complex Gaussian random variable. Under hypothesis H_0, z_p is zero-mean

$$E\{z_p; H_0\} = \frac{1}{\gamma} \mathbf{s}_p^\dagger \mathbf{R}_p^{-1} E\{\mathbf{y}_{i+n_p}\} = 0 \tag{4.50}$$

and it variance is given by

$$\mathrm{var}\{z_p; H_0\} = E\{z_p z_p^\dagger; H_0\} = \frac{1}{\gamma_p^2} \mathbf{s}_p^\dagger \mathbf{R}_p^{-1} E\{\mathbf{y}_{i+n_p} \mathbf{y}_{i+n_p}^\dagger\} \mathbf{R}_p^{-1} \mathbf{s}_p = \frac{1}{\gamma_p^2} \mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{s}_p \tag{4.51}$$

since $E\{\mathbf{y}_{i+n_p}\mathbf{y}_{i+n_p}^\dagger\} = \mathbf{R}_p$. Hence defining

$$\sigma_{0_p}^2 = \frac{1}{\gamma_p^2}\mathbf{s}_p^\dagger\mathbf{R}_p^{-1}\mathbf{s}_p = \frac{\sigma_{\alpha_p}^2\mathbf{s}_p^\dagger\mathbf{R}_p^{-1}\mathbf{s}_p}{\sigma_{\alpha_p}^2\mathbf{s}_p^\dagger\mathbf{R}_p^{-1}\mathbf{s}_p + 1} \tag{4.52}$$

yields

$$H_0 : z_p \sim \mathcal{CN}(0, \sigma_{0_p}^2). \tag{4.53}$$

Similarly, under hypothesis H_1, z_p is also zero-mean

$$E\{z_p; H_1\} = 0 \tag{4.54}$$

and its variance is given by

$$\begin{aligned}\mathrm{var}\{z_p; H_1\} = E\{z_p z_p^\dagger; H_1\} &= \frac{1}{\gamma_p^2}(\mathbf{s}_p^\dagger\mathbf{R}_p^{-1}\mathbf{s}_p + E\{|\alpha_p|^2\}|\mathbf{s}_p^\dagger\mathbf{R}_p^{-1}\mathbf{s}_p|^2) \\ &= \frac{1}{\gamma_p^2}(\mathbf{s}_p^\dagger\mathbf{R}_p^{-1}\mathbf{s}_p + \sigma_{\alpha_p}^2|\mathbf{s}_p^\dagger\mathbf{R}_p^{-1}\mathbf{s}_p|^2) \\ &= \sigma_{\alpha_p}^2\mathbf{s}_p^\dagger\mathbf{R}_p^{-1}\mathbf{s}_p.\end{aligned} \tag{4.55}$$

Defining

$$\sigma_{1_p}^2 = \sigma_{\alpha_p}^2\mathbf{s}_p^\dagger\mathbf{R}_p^{-1}\mathbf{s}_p \tag{4.56}$$

yields

$$H_1 : z_p \sim \mathcal{CN}(0, \sigma_{1_p}^2). \tag{4.57}$$

As z_p is a zero-mean complex Normal distributed random variable, $T_{b,s}(\mathbf{y}_p) = |z_p|^2$ is distributed according to a central chi-squared distribution with 2 degrees of freedom

$$T_{b,s}(\mathbf{y}_p) \sim \begin{cases} \dfrac{\sigma_{0_p}^2}{2}\chi_2^2 & H_0 \\ \dfrac{\sigma_{1_p}^2}{2}\chi_2^2 & H_1. \end{cases} \tag{4.58}$$

Detection is possible because $\sigma_{0_p}^2 < \sigma_{1_p}^2$. Figure 4.2 (a) illustrates the distribution of the test statistic $T_{b,s}(\mathbf{y}_p)$ under both hypotheses.

Setting the threshold to $\eta_{b,s}$, yields

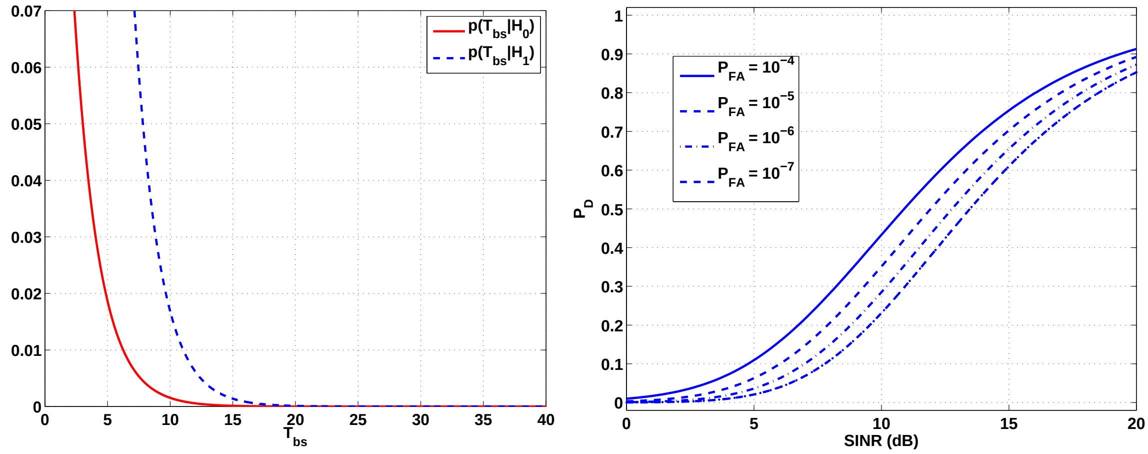
$$P_{\mathrm{FA}} = e^{-\frac{\eta_{b,s}}{\sigma_{0_p}^2}} \tag{4.59}$$

and

$$P_{\mathrm{D}} = e^{-\frac{\eta_{b,s}}{\sigma_{1_p}^2}} \tag{4.60}$$

or, eliminating $\eta_{b,s}$ between these two equations,

$$P_{\mathrm{D}} = P_{\mathrm{FA}}^{\frac{\sigma_{0_p}^2}{\sigma_{1_p}^2}} = P_{\mathrm{FA}}^{\frac{1}{1+\mathrm{SINR}_{p,s}}} \tag{4.61}$$



(a) (b)

Figure 4.2: (a) Plot of the PDF of the test statistic under the two hypotheses and (b) plot of the probability of detection P_{D} as a function of the SINR for different probabilities of false alarm P_{FA}.

where the signal to interference-plus-noise ratio for the Swerling-I target model $\mathrm{SINR}_{p,s}$ is defined as $\mathrm{SINR}_{p,s} = \sigma^2_{\alpha_p}\mathbf{s}_p^\dagger \mathbf{R}_p^{-1}\mathbf{s}_p$. Hence, if the signal-to-noise ratio is zero, one has $P_{\mathrm{D}} = P_{\mathrm{FA}}$ and for a very large $\mathrm{SINR}_{p,s}$, one has $P_{\mathrm{D}} \approx 1$. This is illustrated in Figure 4.2 (b) where the probability of detection is plotted as a function of the SINR for different probabilities of false alarm. Obviously, the probability of detection increases with the SINR and, as the SINR is proportional to the number N_p of channels on the receiver of the bistatic radar p, the probability of detection increases with the number of channels.

4.4.2.3 Discussion

Figure 4.3 shows the detection performance achieved with the detectors presented in the previous section for the considered target models. For relatively high SINR, a target obeying a Swerling-I model (fluctuating complex amplitude) is more difficult to detect than a target obeying a Marcum model (deterministic complex amplitude). The fact that fluctuating targets are more difficult to detect is well known and can be explained by the fact that the amplitude of the returned echo fluctuates. Therefore, for very weak echoes, it is possible that the detection statistic does not exceed the detection threshold. However, for smaller SINR, the opposite occurs: a fluctuating target is easier to detect than a steady target. Indeed, if the target is fluctuating, there is a non-zero probability that a high enough amplitude of the target echo exceeds the the detection threshold and that a detection occurs.

4.4.3 Multistatic scenarios

Let us consider multistatic scenarios, i.e., scenarios for which the detection is based on the combination of measurements produced by several bistatic radars.

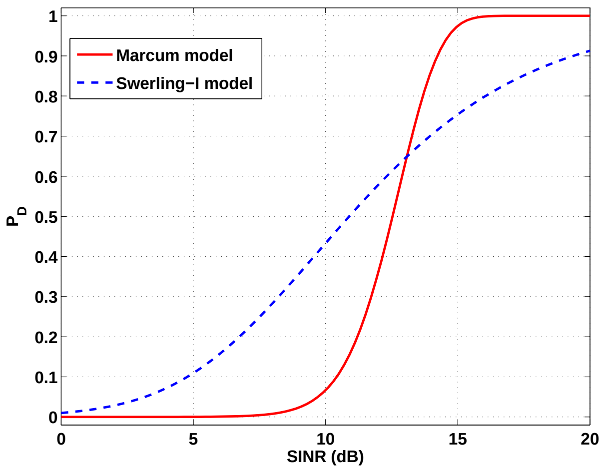


Figure 4.3: Comparison of the detection performance for the considered target models ($P_{\text{FA}} = 10^{-4}$).

4.4.3.1 Marcum target model

Let us define

$$z_p = \frac{1}{\gamma_p} \mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{y}_p \tag{4.62}$$

where $\gamma_p = \sqrt{\mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{s}_p}$. Section 4.4.2.1, yields

$$z_p \sim \begin{cases} \mathcal{CN}(0, 1) & H_0 \\ \mathcal{CN}(\gamma_p \alpha_p, 1) & H_1 \end{cases} \tag{4.63}$$

and from (4.27),

$$T_{m,m}(\mathbf{y}) = \sum_{p=1}^{P} |z_p|^2 = \sum_{p=1}^{P} \text{Re}(z_p)^2 + \text{Im}(z_p)^2. \tag{4.64}$$

Under H_0, the test statistic is thus the sum of $2P$ squares of zero-mean normal random variables with a variance equal to $\frac{1}{2}$. The corresponding distribution is a chi-squared with $2P$ degrees of freedom. Under H_1, the test statistic is the sum of $2P$ squares of normal random variables with non-zero mean. The corresponding distribution is a non-central chi-squared with $2P$ degrees of freedom and non-centrality parameter $\mu = 2\sum_{p=1}^{P} \gamma_p^2 |\alpha_p|^2 = 2\sum_{p=1}^{P} \text{SINR}_{p,m}$:

$$T_{m,m}(\mathbf{y}) \sim \begin{cases} \frac{1}{2}\chi^2_{2P} & H_0 \\ \frac{1}{2}\chi'^2_{2P}(\mu) & H_1. \end{cases} \tag{4.65}$$

The threshold $\eta_{m,m}$ for a given P_{FA} is obtained by computing

$$\eta_{m,m} = \frac{1}{2} Q^{-1}_{\chi^2_{2P}}(P_{\text{FA}}) \tag{4.66}$$

and the probability of detection P_{D} is

$$P_{\mathrm{D}} = Q_{\chi'^2_{2P}(\mu)}(2\eta_{m,m}). \tag{4.67}$$

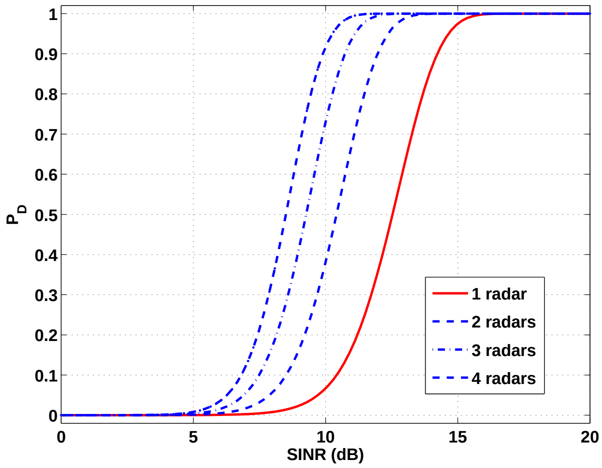


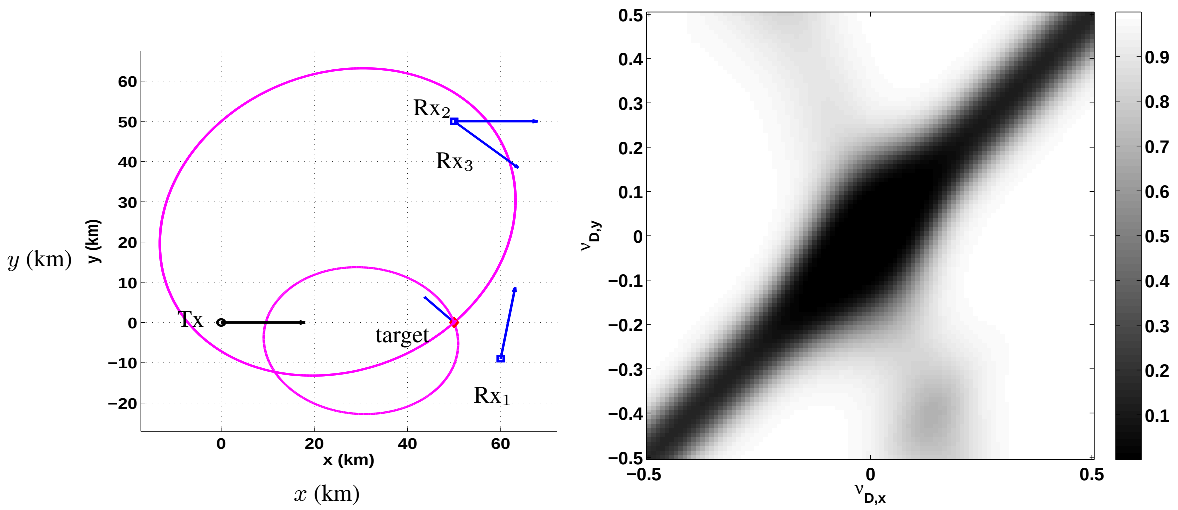
Figure 4.4: Influence of the number of (bistatic) radars P on the probability of detection in the case of Marcum target model ($P_{\mathrm{FA}} = 10^{-4}$)

Figure 4.4 shows the evolution of the probability of detection as a function of the SINR, assumed identical for all radars in order to simplify the graph, for varying numbers of contributing bistatic radars. For a given SINR, the probability of detection increases with the number of contributing radars. The combination of the contributions of the different bistatic radars is similar to an incoherent summation.

The SINR of individual bistatic radars generally varies according to the spectral signature of the clutter and the spatio-temporal location of the target, which will have an impact on the detection performance. By way of illustration, let us consider a particular multistatic scenario as represented in Figure 4.5 (a), which depicts one transmitter (Tx) and three receive platforms (Rx_1, Rx_2, and Rx_3). The arrows indicate the velocity of the target and of the radar platforms. The ellipses depict the bistatic isoranges on the ground for each bistatic radar pair. The clutter patches inducing interferences competing with the target return are located along these isoranges. The transmit and receive platforms are airborne, which explains why they do not appear at the focal point of the ellipses. The platforms Rx_2 and Rx_3 are nearly collocated but have different speeds. The probability of detection for a probability of false alarm equal to 10^{-4} and for different target velocities[5] is depicted in Figure 4.5 (b). The magnitude $|\alpha_p|$ of the target echo signal is taken identical for all platforms and the setup is such that the maximum SINR (i.e., the SINR when the hypothesized target has a high Doppler frequency) is about 12.7dB[6] for a single radar.

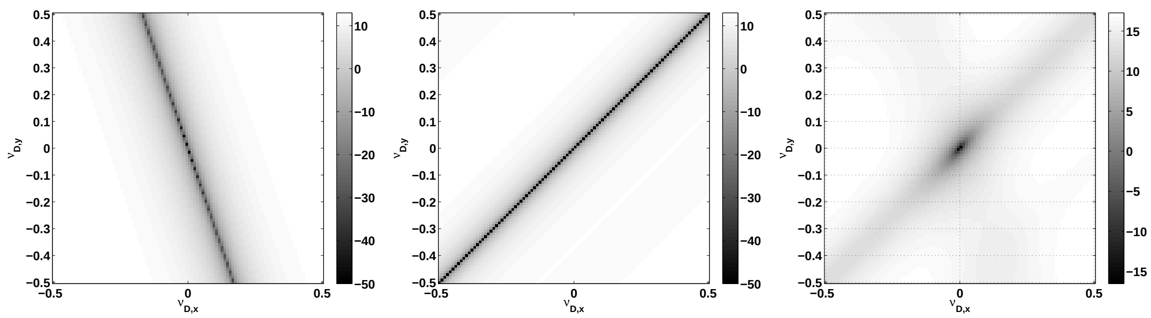
[5] The reduced target velocity is defined by $\nu_{Dx} = \frac{v_x}{f_p \lambda}$ where v_x is the velocity along the x-axis and similarly for ν_{Dy}

[6] This value is taken such that the loss in detection performance if the target has zero range rate with respect to one single radar is visible on the graph. It is clear that if the target echo magnitude is higher, the detection probability can be close to unity for all velocity vectors with non-zero magnitude.



(a) (b)

Figure 4.5: (a) The considered multistatic scenario and (b) probability of detection as a function of the reduced target velocity vector (ν_{Dx}, ν_{Dy}) in the case of a target corresponding to the Marcum target model with maximum SINR of 12.66dB ($P_{\mathrm{FA}} = 10^{-4}$).



(a) (b) (c)

Figure 4.6: (a) $\mathrm{SINR}_{1,m}$ (b) $\mathrm{SINR}_{2,m}$ (c) $\mu/2 = \sum_{p=1}^{3} \mathrm{SINR}_{p,m}$ for the scenario of Figure 4.5.

The probability of detection at zero velocity is of course very small, due to the presence of the clutter and the corresponding low SINR. Monostatic radars are only sensitive to the radial velocity component, which means that targets traveling perpendicular to the line of sight of the radar are indistinguishable from the clutter. There is a similar issue with bistatic radars [174] for which targets having a velocity tangential to the isorange have zero range-rate and hence do not exhibit a bistatic Doppler, which explains the loss of detection performance along the "diagonal" $\nu_{Dx} = \nu_{Dy}$ in Figure 4.5 (b), corresponding to the velocities for which the target appears "static" to the bistatic radars Tx-Rx$_2$ and Tx-Rx$_3$. Indeed, these velocities are tangential to the isorange corresponding to the radars Tx-Rx$_2$ and Tx-Rx$_3$. For these velocities, the SINR corresponding to these two bistatic radars is effectively very small as can be seen on Figure 4.6 (b) ($\mathrm{SINR}_{3,m}$, visually similar to $\mathrm{SINR}_{2,m}$, is therefore not presented) which explains the decrease of probability of detection. The probability of detection is however not equal to zero thanks to the radar Tx-Rx$_1$ which observes the target with a different geometry, for which the SINR is large. This is illustrated in Figure 4.6 (c) in which the non-centrality parameter is depicted. Indeed, the notch in SINR due to the target, appearing as static to certain radars only, essentially vanishes. There is another direction in which there is a loss in detection performance, corresponding to velocities tangential to the isorange of bistatic radar Tx-Rx$_1$. For those velocities, the contribution of bistatic radar Tx-Rx$_1$ is small as can be seen on Figure 4.6 (a).

4.4.3.2 Swerling-I target model

Let us define

$$z_p = \frac{1}{\gamma_p} \mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{y}_p \tag{4.68}$$

where $\gamma_p = \sqrt{\frac{1}{\sigma_{\alpha_p}^2} + \mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{s}_p}$. Section 4.4.2.2, yields

$$z_p \sim \begin{cases} \mathcal{CN}(0, \sigma_{0_p}^2) & H_0 \\ \mathcal{CN}(0, \sigma_{1_p}^2) & H_1 \end{cases} \tag{4.69}$$

hence from (4.30)

$$T_{m,s}(\mathbf{y}) = \sum_{p=1}^{P} |z_p|^2 = \sum_{p=1}^{P} \mathrm{Re}(z_p)^2 + \mathrm{Im}(z_p)^2. \tag{4.70}$$

The test statistic is thus the sum of the squares of zero-mean normal distributed random variables, each having different variances.

To compute the probability of detection, the knowledge of the PDF of $T_{m,s}(\mathbf{y})$ is needed. The PDF of the sum of squares of normal distributed random variables with different variances is not known. In order to further pursue the analysis, let us assume that $\sigma_{0_p}^2 = \sigma_0^2$, $\forall p$ and $\sigma_{1_p}^2 = \sigma_1^2$, $\forall p$. This assumption implies that the $\mathrm{SINR}_{p,s} = \sigma_{\alpha_p}^2 \mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{s}_p$ is the same for all bistatic radars, which is equivalent to assuming that 1) the target RCS is the same for all bistatic radars and 2) the interference-plus-noise power is also the same for each radar. Although this assumption is very restrictive, the goal is to analyze the effect of the angular diversity and this effect is still adequately modeled despite this assumption. In these conditions, the test statistic is distributed according to a scaled central chi-squared distribution with $2P$ degrees of freedom:

$$T_{m,s}(\mathbf{y}) \sim \begin{cases} \frac{\sigma_0^2}{2} \chi_{2P}^2 & H_0 \\ \frac{\sigma_1^2}{2} \chi_{2P}^2 & H_1 \end{cases} \tag{4.71}$$

The threshold $\eta_{m,s}$ for a given P_{FA} is obtained by computing

$$\eta_{m,s} = \frac{\sigma_0^2}{2} Q_{\chi_{2P}^2}^{-1}(P_{\mathrm{FA}}) \tag{4.72}$$

and the probability of detection P_{D} is given by

$$P_{\mathrm{D}} = Q_{\chi^2_{2P}}\left(\frac{2}{\sigma_1^2}\eta_{m,s}\right). \tag{4.73}$$

Figure 4.7 shows the evolution of the probability of detection as a function of the SINR and of

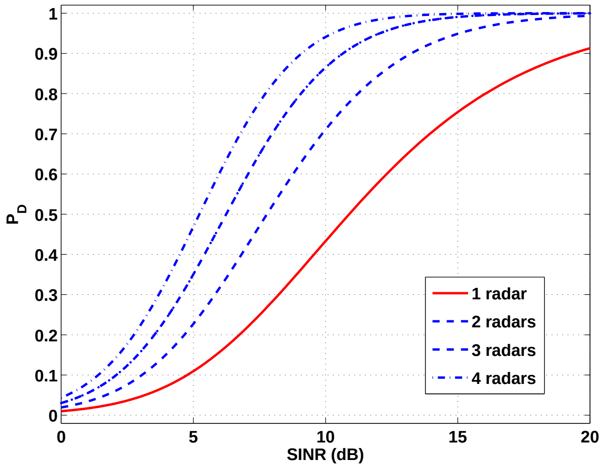


Figure 4.7: Influence of the number of (bistatic) radars P on the probability of detection for Swerling-I targets ($P_{\mathrm{FA}} = 10^{-4}$).

the number of contributing bistatic radars. As can be seen, for a given SINR, the probability of detection increases with the number of contributing radars. This is a manifestation of the diversity gain and can be explained as follows: even if the complex amplitude of the target echo received by one radar is small, there is however some chance to get a detection thanks to a large enough target echo from another radar. The larger the number of contributing radars, the more likely the reception by one of the radars of an echo with a large amplitude.

In order to further assess the performance of a multistatic radar in the case of varying SINR in a less restrictive scenario, let us consider the *deflection coefficient* [53, 82, 169] defined as

$$d^2 = \frac{|E\{T; H_1\} - E\{T; H_0\}|^2}{\frac{1}{2}(\mathrm{var}\{T; H_1\} + \mathrm{var}\{T; H_0\})} \tag{4.74}$$

where T is the considered test statistic. The deflection coefficient is a measure of the distance between two distributions. If T is Gaussian under both hypotheses, which is not the case here, the deflection coefficient entirely characterizes the detection performance [82]. However, the deflection coefficient still provides an indication of the detection performance [43]. As $|z_p|^2$ is distributed according to

$$|z_p|^2 \sim \begin{cases} \dfrac{\sigma_{0_p}^2}{2}\chi_2^2 & H_0 \\ \dfrac{\sigma_{1_p}^2}{2}\chi_2^2 & H_1, \end{cases} \tag{4.75}$$

the properties of the central χ^2 distribution lead to

$$E\{|z_p|^2; H_0\} = \sigma^2_{0_p} \tag{4.76}$$
$$E\{|z_p|^2; H_1\} = \sigma^2_{1_p} \tag{4.77}$$
$$\mathrm{var}\{|z_p|^2; H_0\} = \sigma^4_{0_p} \tag{4.78}$$
$$\mathrm{var}\{|z_p|^2; H_1\} = \sigma^4_{1_p}. \tag{4.79}$$

The independence of $|z_p|^2$ for different p yields

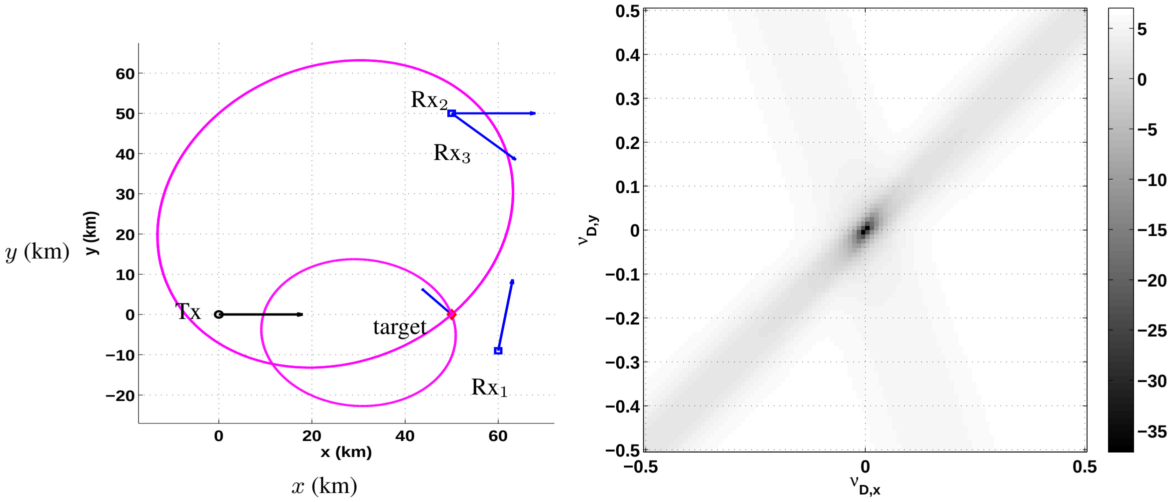
$$E\{T_{m,s}; H_0\} = \sum_{p=1}^{P} \sigma^2_{0_p} \tag{4.80}$$
$$E\{T_{m,s}; H_1\} = \sum_{p=1}^{P} \sigma^2_{1_p} \tag{4.81}$$
$$\mathrm{var}\{T_{m,s}; H_0\} = \sum_{p=1}^{P} \sigma^4_{0_p} \tag{4.82}$$
$$\mathrm{var}\{T_{m,s}; H_1\} = \sum_{p=1}^{P} \sigma^4_{1_p} \tag{4.83}$$

hence

$$E\{T_{m,s}; H_1\} - E\{T_{m,s}; H_0\} = \sum_{p=1}^{P} \sigma^2_{1_p} - \sigma^2_{0_p} = \sum_{p=1}^{P} \frac{|\sigma^2_{\alpha_p} \mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{s}_p|^2}{\sigma^2_{\alpha_p} \mathbf{s}_p^\dagger \mathbf{R}_p^{-1} \mathbf{s}_p + 1} \tag{4.84}$$
$$\mathrm{var}\{T_{m,s}; H_1\} + \mathrm{var}\{T_{m,s}; H_0\} = \sum_{p=1}^{P} \sigma^4_{1_p} + \sigma^4_{0_p}. \tag{4.85}$$

As illustration, let us consider the same multistatic scenario as in previous section and depicted again in Figure 4.8 (a) for convenience. The deflection coefficient obtained for different target velocities is depicted in Figure 4.8 (b). The target scattering coefficient σ_{α_p} is assumed identical for all platforms and the setup is such that the maximum SINR for a single radar is about 10dB. As can be seen, there is a huge loss in detection performance (small deflection coefficient) at zero velocity, corresponding to the clutter notch. The loss in detection performance along the diagonal $\nu_{Dx} = \nu_{Dy}$ is clearly visible and is explained in the same way as in the previous section (zero target range-rate[7] with respect to the bistatic radars Tx-Rx$_2$ and Tx-Rx$_3$) while the loss in deflection coefficient due to the zero target range-rate with respect to the bistatic radar Tx-Rx$_3$ is barely visible.

[7] The bistatic range-rate is considered.



(a) (b)

Figure 4.8: (a) Considered multistatic scenario and (b) deflection coefficient (in dB) as a function of the reduced target velocity vector.

4.4.3.3 Discussion

From the previous discussion it is clear that an increase of the number of contributing bistatic radars results in an increase of the probability of detection as illustrated in Figures 4.4 and 4.7. However, this comparison may be misleading as the cost incurred for this increase in performance, i.e., the total number of receive elements, is not taken into account. More interesting is to know, given a total number of receive elements N, how to combine them to yield the best detection performance. A higher number of receive elements per radar increases the SINR of that particular radar and hence increases the detection performance of that radar. However, it is expected that having more radars also increases the detection performance. Thus, the question is to which extent an increase of the SINR of the individual radars can actually compensate for the loss of the angular diversity gain. Figure 4.9 illustrates the expected detection performance if the total number N of receive elements is kept constant. For this comparison, the SINR is kept identical for all radars, which corresponds to a case for which the target has a non-zero (and large) range-rate with respect to all bistatic radars.

In the case of a Marcum target model (Figure 4.9 (a)), there is actually a *decrease* in detection performance if the number of available receive elements is distributed across the different radars. In this case indeed, there is no diversity gain. On the contrary, if all available receive elements are concentrated in one single radar, a higher coherent integration gain is secured, hence a higher SINR for that radar.

The situation is however totally different if a Swerling-I target model is considered (Figure 4.9 (b)). Clearly, for relatively high SINR, it is advantageous to have several radars, as the angular diversity is favored at the cost of the SINR of individual radars. The contrary is true for low SINR. Note that conclusion differs from the one of Section 4.4.3.2 in which it was found always advantageous to increase the number of radars. If the total number of receive elements is

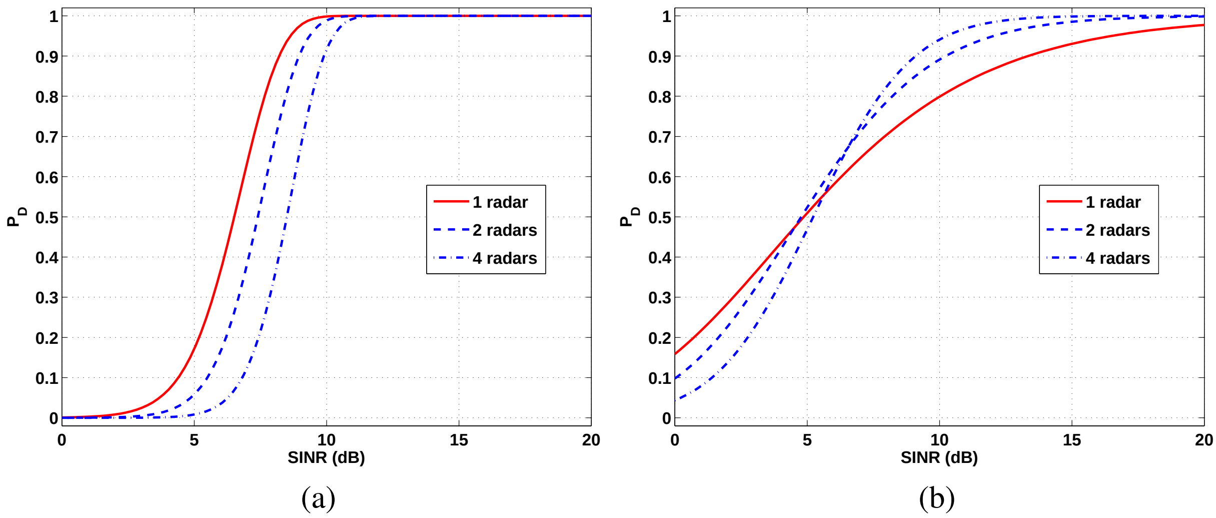


Figure 4.9: Comparison of the detection performance of different scenarios with a constant total number of receive elements (a) Marcum target model and (b) Swerling-I target model ($P_{\mathrm{FA}} = 10^{-4}$).

kept constant, at low SINR, the angular diversity gain does not make up for the low SINR.

4.5 Realizable detectors

Generally speaking, the PDF of the test statistic T depends on the interference-plus-noise covariance matrix $\mathbf{R}_p$ and the target steering vector $\mathbf{s}_p$. The expressions of the test statistic obtained in Section 4.3 assumed that these parameters are known. In practice, this is not the case.

These parameters are called nuisance parameters [82] as their actual value is not of direct interest for our detection problem. It would be desirable to have a test which would not require the knowledge of these nuisance parameters. This kind of test is called uniformly most powerful. However such a test exists if and only if the likelihood ratio test can be completely defined without knowledge of the nuisance parameters [169]. This is clearly not the case here. A solution consists in going back to the generalized likelihood ratio, described in Section 4.3.2.

The next sections discuss these issues.

4.5.1 Unknown interference-plus-noise covariance matrix

The generalized likelihood is equivalent to replacing the clairvoyant covariance matrix $\mathbf{R}_p$ by its maximum likelihood estimate $\hat{\mathbf{R}}_p$ in the expression of the test statistic. In other words, the weight vector of (4.22) becomes

$$\hat{\mathbf{w}}_p = \hat{k}\hat{\mathbf{R}}_p^{-1}\mathbf{s}_p. \tag{4.86}$$

The two next sections discuss how the hypothesis tests are transformed by using $\hat{\mathbf{R}}_p$ instead of $\mathbf{R}_p$ and how to characterize their performance. It is clear that the filter $\hat{\mathbf{w}}_p$ is random as $\hat{\mathbf{R}}_p$ is itself random. However, in order to characterize the effect of using the filter $\hat{\mathbf{w}}_p$ instead of the optimal filter $\mathbf{w}_p$, the following sections use a particular realization of $\hat{\mathbf{w}}_p$, considered as deterministic.

4.5.1.1 Marcum target model

The generalized likelihood ratio test (4.12) becomes

$$\frac{|\mathbf{s}_p^\dagger \hat{\mathbf{R}}_p^{-1} \mathbf{y}_p|^2}{\mathbf{s}_p^\dagger \hat{\mathbf{R}}_p^{-1} \mathbf{s}_p} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{b,m} \tag{4.87}$$

also known as the modified sample matrix inversion (MSMI) [114] test statistic. It can be shown [150] that this test retains its CFAR property. As far as the detection performance is concerned, the same reasoning as in Section 4.4.2.1 holds, by defining

$$\hat{z}_p = \frac{1}{\hat{\gamma}_p} \mathbf{s}_p^\dagger \hat{\mathbf{R}}_p^{-1} \mathbf{y}_p \tag{4.88}$$

with $\hat{\gamma}_p = \sqrt{\mathbf{s}_p^\dagger \hat{\mathbf{R}}_p^{-1} \mathbf{s}_p} = \sqrt{\frac{1}{\hat{k}} \hat{\mathbf{w}}_p^\dagger \mathbf{s}_p}$. It is easy to show that

$$E\{\hat{z}_p; H_0\} = 0 \tag{4.89}$$

$$E\{\hat{z}_p; H_1\} = \hat{\gamma}_p \alpha_p \tag{4.90}$$

$$\mathrm{var}\{\hat{z}_p; H_0\} = \mathrm{var}\{\hat{z}_p; H_1\} = \frac{\mathbf{s}_p^\dagger \hat{\mathbf{R}}_p^{-1} \mathbf{R}_p \hat{\mathbf{R}}_p^{-1} \mathbf{s}_p}{\hat{\gamma}_p^2} = \frac{\hat{\mathbf{w}}_p^\dagger \mathbf{R}_p \hat{\mathbf{w}}_p}{\hat{k} \hat{\mathbf{w}}_p^\dagger \mathbf{s}_p} = \sigma_{m_p}^2, \tag{4.91}$$

which defines σ_{m_p}, hence $|\hat{z}_p|^2$ has the same PDF as $|z_p|^2$ up to a scale factor and a different non-centrality parameter

$$|\hat{z}_p|^2 \sim \begin{cases} \dfrac{\sigma_{m_p}^2}{2} \chi_2^2 & H_0 \\ \dfrac{\sigma_{m_p}^2}{2} \chi_2'^2 \left(2 \dfrac{\hat{\gamma}_p^2 |\alpha_p|^2}{\sigma_{m_p}^2} \right) & H_1. \end{cases} \tag{4.92}$$

As $\hat{\mathbf{w}}_p^\dagger \mathbf{s}_p = \hat{k} \mathbf{s}_p^\dagger \hat{\mathbf{R}}_p^{-1} \mathbf{s}_p$ is a quadratic form hence real, the non-centrality parameter of the $\chi_2'^2$ distribution can be rewritten as

$$2 \frac{\hat{\gamma}_p^2 |\alpha_p|^2}{\sigma_{m_p}^2} = 2 \frac{|\hat{\mathbf{w}}_p^\dagger \mathbf{s}_p|^2}{\hat{\mathbf{w}}_p^\dagger \mathbf{R}_p \hat{\mathbf{w}}_p} |\alpha_p|^2 = 2 \widehat{\mathrm{SINR}}_{p,m} \tag{4.93}$$

where $\widehat{\mathrm{SINR}}_{p,m}$ is the signal to interference-plus-noise ratio at the output of the filter

$$\hat{\mathbf{w}}_p = \hat{k} \hat{\mathbf{R}}_p^{-1} \mathbf{s}_p. \tag{4.94}$$

As discussed in Section 4.4.2.1, the probability of detection increases with the non-centrality parameters, i.e., with the signal to interference-plus-noise ratio $\widehat{\mathrm{SINR}}_{p,m}$.

4.5.1.2 Swerling-I target model

The likelihood ratio test (4.18) becomes

$$\frac{|\mathbf{s}_p^\dagger \hat{\mathbf{R}}_p^{-1} \mathbf{y}_p|^2}{\frac{1}{\sigma_{\alpha_p}^2} + \mathbf{s}_p^\dagger \hat{\mathbf{R}}_p^{-1} \mathbf{s}_p} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{b,s}. \tag{4.95}$$

Following the same developments as in Section 4.4.2.2 and defining

$$\hat{z}_p = \frac{1}{\hat{\gamma}_p} \mathbf{s}_p^\dagger \hat{\mathbf{R}}_p^{-1} \mathbf{y}_p \tag{4.96}$$

where $\hat{\gamma}_p = \sqrt{\frac{1}{\sigma_{\alpha_p}^2} + \mathbf{s}_p^\dagger \hat{\mathbf{R}}_p^{-1} \mathbf{s}_p}$, yields easily

$$|\hat{z}_p|^2 \sim \begin{cases} \dfrac{\hat{\sigma}_{0_p}^2}{2} \chi_2^2 & H_0 \\ \dfrac{\hat{\sigma}_{1_p}^2}{2} \chi_2^2 & H_1 \end{cases} \tag{4.97}$$

where

$$\hat{\sigma}_{0_p}^2 = \frac{1}{\hat{\gamma}_p^2} \mathbf{s}_p^\dagger \hat{\mathbf{R}}_p^{-1} \mathbf{R}_p \hat{\mathbf{R}}_p^{-1} \mathbf{s}_p = \frac{1}{\hat{k}^2 \hat{\gamma}_p^2} \hat{\mathbf{w}}_p^\dagger \mathbf{R}_p \hat{\mathbf{w}}_p \tag{4.98}$$

$$\hat{\sigma}_{1_p}^2 = \hat{\sigma}_{0_p}^2 + \frac{1}{\hat{\gamma}_p^2} \sigma_{\alpha_p}^2 |\mathbf{s}_p^\dagger \hat{\mathbf{R}}_p^{-1} \mathbf{s}_p|^2 = \hat{\sigma}_{0_p}^2 + \frac{1}{\hat{k}^2 \hat{\gamma}_p^2} \sigma_{\alpha_p}^2 |\hat{\mathbf{w}}_p^\dagger \mathbf{s}_p|^2 \tag{4.99}$$

$$\tag{4.100}$$

and P_{D} is still given by

$$P_{\mathrm{D}} = P_{\mathrm{FA}}^{\frac{\hat{\sigma}_{0p}^2}{\hat{\sigma}_{1p}^2}} = P_{\mathrm{FA}}^{\frac{1}{1+\hat{\mathrm{SINR}}_{p,s}}} \tag{4.101}$$

where

$$\hat{\mathrm{SINR}}_{p,s} = \sigma_{\alpha_p}^2 \frac{|\mathbf{s}_p^\dagger \hat{\mathbf{R}}_p^{-1} \mathbf{s}_p|^2}{\mathbf{s}_p^\dagger \hat{\mathbf{R}}_p^{-1} \mathbf{R}_p \hat{\mathbf{R}}_p^{-1} \mathbf{s}_p} = \sigma_{\alpha_p}^2 \frac{|\hat{\mathbf{w}}_p^\dagger \mathbf{s}_p|^2}{\hat{\mathbf{w}}_p^\dagger \mathbf{R}_p \hat{\mathbf{w}}_p} \tag{4.102}$$

is the SINR at the output of the filter

$$\hat{\mathbf{w}}_p = \hat{k} \hat{\mathbf{R}}_p^{-1} \mathbf{s}_p. \tag{4.103}$$

Thus the performance of this detector only depends on the SINR achieved at the output of the filter $\hat{\mathbf{w}}_p$.

4.5.1.3 Performance measure

The fact that the detection performance increases with the signal to interference-plus-noise ratio $\hat{\text{SINR}}_p$ justifies the use of the SINR loss as a measure of the performance of the filter $\hat{\mathbf{w}}_p$. The SINR loss is defined as

$$\text{SINR}_{\text{loss}p} = \frac{\hat{\text{SINR}}_p}{\text{SNR}_p} = \frac{\hat{\text{SINR}}_p}{\text{SINR}_p}\frac{\text{SINR}_p}{\text{SNR}_p} \tag{4.104}$$

where SNR_p is the signal to thermal noise ratio of bistatic radar p and SINR_p the SINR achieved at the output of the optimum filter $\mathbf{w}_p$.

For the Marcum target model, $\text{SNR}_{p,m} = \frac{|\alpha_p|^2}{\sigma_n^2}\mathbf{s}^\dagger\mathbf{s}$ and $\text{SINR}_{p,m} = |\alpha_p|^2\mathbf{s}_p^\dagger\mathbf{R}_p^{-1}\mathbf{s}_p$. For the Swerling-I target model, $\text{SNR}_{p,s} = \frac{\sigma_{\alpha_p}^2}{\sigma_n^2}\mathbf{s}^\dagger\mathbf{s}$ and $\text{SINR}_{p,s} = \sigma_{\alpha_p}^2\mathbf{s}_p^\dagger\mathbf{R}_p^{-1}\mathbf{s}_p$. In both cases σ_n^2 is the thermal noise variance.

The first factor of $\text{SINR}_{\text{loss}p}$ represents the loss due the use of an estimated interference-plus-noise covariance matrix $\hat{\mathbf{R}}_p$ instead of the clairvoyant interference-plus-noise covariance matrix $\mathbf{R}_p$ and the second factor represents the loss due to the presence of the interference if the optimum filter $\mathbf{w}_p$ is used. The $\text{SINR}_{\text{loss}p}$ is commonly used to characterize the quality of the estimate $\hat{\mathbf{R}}_p$ and more generally, the quality of filter $\hat{\mathbf{w}}_p$.

However, the use of the $\text{SINR}_{\text{loss}p}$ to characterize the performance of the filter does not extend to the general multistatic case due to the lack of known PDF for the test statistic. However, individually maximizing the detection performance of each bistatic radar, i.e., independently optimizing the filter for each radar, intuitively seems reasonable.

4.5.2 Unknown steering vector parameters

To compute the optimum filter $\mathbf{w}_p = k\mathbf{R}_p^{-1}\mathbf{s}_p$, the knowledge of the target steering vector

$$\mathbf{s}_p = \mathbf{s}_p(\boldsymbol{\theta}_p(\vec{x}_t)) = \mathbf{s}_p(\vec{\nu}_{p_s}, \nu_{p_D}) \tag{4.105}$$

is assumed.

However, the position and velocity of the target is typically unknown hence the spatial frequency vector $\vec{\nu}_{p_s}$, the Doppler frequency ν_{p_D}, and also the range r_p where to perform the filtering are unknown parameters of the likelihood ratio. The commonly accepted approach [82] consists in using the generalized likelihood ratio

$$\Lambda_{g_p} = \frac{\max\limits_{\vec{\nu}_{p_s},\nu_{p_D},r_p} p(\mathbf{y}_p; \vec{\nu}_{p_s}, \nu_{p_D}, r_p, H_1)}{p(\mathbf{y}_p; H_0)} \tag{4.106}$$

which is a suboptimal solution. In practice, the likelihood ratio is maximized over the unknown parameters. This maximization is usually [82] performed by sampling the parameter space $\{\vec{\nu}_{p_s}, \nu_{p_D}, r_p\}$ or equivalently, by sampling the spatial location of the target $\vec{x}_t$ and its velocity $\dot{\vec{x}}_t$.

Of course, the hypothesized steering vector $\hat{\mathbf{s}}_p$ does not exactly correspond to the actual target steering vector $\mathbf{s}_p$. This so-called signal mismatch [105] causes an additional loss in the signal to interference-plus-noise ratio. Expressions for this loss are given in [105, 150]. Channel mismatch occurring as the receive elements differ from each other in phase or in gain [60] is another cause of signal mismatch. In the latter case, the steering vectors used to process the data differ from the physical steering vectors. Other phenomena induce differences between the physical steering vectors and the modeled ones. Generally, the use of measured steering vectors is required in real applications [5, 153].

As the estimation of the interference-plus-noise covariance matrix constitutes our major concern, let us assume that no signal mismatch occurs.

4.6 Centralized versus decentralized detection

The previous sections have presented optimum detection methods which require to send the test statistic computed by each bistatic radar to a central site which performs final combination of the test statistic and, finally, the detection by thresholding the resulting test statistic. This process, by its very nature, is a centralized detection scheme.

Other methods have been proposed [32, 53, 160] in which the detection is performed individually for each radar and is followed by a more simple combination in the central site. A decentralized processing offers the advantage of reduced bandwidth requirements between the individual radars and the central site, as only the coordinates of the detections need to be transmitted, if relatively few detections occurs. [160] shows that the test statistic of the optimum decentralized detector is a weighted sum of the (binary) detection of each radar. The detection itself, under some conditions, can be implemented by a simple logical sum, i.e., a binary OR operation [32, 160]. Although the decentralized detection scheme is suboptimal, the performance loss is limited to about 3dB if the number of contributing radars is reasonable (< 10) and for high SINR's (> 16dB) [32].

Chapter 5

Clutter signal covariance matrix

5.1 Introduction

The computation of the optimum detector detailed in the previous chapter requires the clutter signal covariance matrix. Although the estimation of this matrix from the radar data is discussed in Chapter 7, it is appropriate to analyze how this matrix varies for different radar configuration parameters, particularly for different ranges to the target and/or for different antenna array shapes.

The analysis of the clutter covariance matrix is performed in the spectral domain. The power spectrum of a continuous space-time random signal is particularized to the power spectrum of the clutter signal. In order to separate the effects due to the radar configuration (positions and speeds of the platforms) from those due to the spatial and temporal sampling (antenna element positions), a system able to measure the signal at any position in time and in space, which is equivalent to assuming that the system has an infinite resolution and that no aliasing occurs, is first considered. When the system has an aliasing-free infinite resolution, both in space and in time, the power spectrum of the clutter proves to be non-zero only along a curve, called the clutter power spectrum locus, in the 4-dimensional frequency space. The link between this curve in 4 dimensions and the curve obtained by considering a linear array (with infinite spatial resolution) is then shown. The behavior of this curve as a function of the radar configuration parameters and the range induces important properties of the covariance matrix such as range-dependence.

The covariance matrix is obtained by sampling the covariance function of the continuous space-time signal. The relation between the clutter power spectrum locus and estimates of the power spectrum of the covariance matrix is deduced next to demonstrate the influence of the antenna array layout (linear, circular, ...).

5.2 Power spectrum relations

5.2.1 Introduction

This section considers continuous spatio-temporal quantities. An infinite resolution is assumed, which means that continuous spatial antennas with infinite extent, thus with an infinite spatial resolution are considered, as well as a continuous wave with an infinitely long observation interval, thus with an infinite Doppler resolution.

5.2.2 Plane wave

A plane wave with a complex amplitude α and an angular frequency ω arriving from direction (θ, ϕ), as illustrated in Figure 5.1, creates the space-time signal

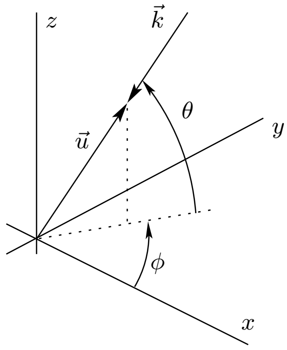


Figure 5.1: Direction of arrival of a plane wave.

$$g'(\vec{r}, t) = \alpha e^{j(\omega t - \vec{k}\cdot\vec{r})} \tag{5.1}$$

where the prime indicates that the signal is at the carrier frequency (rather than at baseband) and where

$$\vec{k} = -\frac{2\pi}{\lambda}\begin{pmatrix}\cos\theta\cos\phi \\ \cos\theta\sin\phi \\ \sin\theta\end{pmatrix} \tag{5.2}$$

is the wave vector $\vec{k} = (k_x, k_y, k_z)$ [78]. This signal is defined in the 4-dimensional space with the 3 spatial coordinates $\vec{r} = (x, y, z)$ and the time t. Let us denote $\vec{u}$ the direction of arrival of the plane wave

$$\vec{u} = \begin{pmatrix}\cos\theta\cos\phi \\ \cos\theta\sin\phi \\ \sin\theta\end{pmatrix} \tag{5.3}$$

and hence

$$\vec{k} = -\frac{2\pi}{\lambda}\vec{u} = -\frac{\omega}{c}\vec{u}. \tag{5.4}$$

Thus (5.1) can be rewritten as

$$g'(\vec{r}, t) = \alpha\; e^{j2\pi(ft + \vec{u}\cdot\frac{\vec{r}}{\lambda})}. \tag{5.5}$$

where $2\pi f = \omega$ is the angular frequency of the signal.

5.2.3 Power spectrum

First, let us consider a continuous space-time random signal $g'(\vec{r}, t)$ with zero mean. If the signal is wide-sense stationary in space and time, its correlation function takes the form

$$r'(\vec{\Delta r}, \Delta t) = E\{g'(\vec{r}, t) g'^{*}(\vec{r} - \vec{\Delta r}, t - \Delta t)\}. \tag{5.6}$$

The 4-dimensional power spectrum of the signal $g'(\vec{r}, t)$ is defined as the Fourier transform of its covariance function [26, 78]

$$P'(\vec{k}, \omega) = \iiiint_{-\infty}^{+\infty} r'(\vec{\Delta r}, \Delta t) e^{-j(\omega \Delta t - \vec{k}\cdot\vec{\Delta r})}\; d\vec{\Delta r}\; d\Delta t. \tag{5.7}$$

This power spectrum can be interpreted as the energy of the plane waves with angular frequency ω and arriving from direction $\vec{k}$ [26, 78].

In the case of a modulated signal one has

$$g'(\vec{r}, t) = e^{j\omega_c t} g(\vec{r}, t) \tag{5.8}$$

where ω_c is the carrier angular frequency, g is the baseband signal, and g' is the modulated signal at the carrier frequency. As discussed in Section 3.3.2, $\omega = \omega_c + \omega_D$ where ω_D is the Doppler angular frequency. The covariance function of g' becomes

$$r'(\vec{\Delta r}, \Delta t) = e^{j\omega_c \Delta t} r(\vec{\Delta r}, \Delta t) \tag{5.9}$$

where $r(\vec{\Delta r}, \Delta t)$ is the covariance function of g with

$$r(\vec{\Delta r}, \Delta t) = E\{g(\vec{r}, t) g^{*}(\vec{r} - \vec{\Delta r}, t - \Delta t)\}. \tag{5.10}$$

The power spectrum of r' is in this case

$$\begin{aligned} P'(\vec{k}, \omega) &= \iiiint_{-\infty}^{+\infty} e^{j\omega_c \Delta t}\; r(\vec{\Delta r}, \Delta t) e^{-j(\omega \Delta t - \vec{k}\cdot\vec{\Delta r})}\; d\vec{\Delta r}\; d\Delta t & (5.11)\\ &= \iiiint_{-\infty}^{+\infty} r(\vec{\Delta r}, \Delta t) e^{-j(\omega_D \Delta t - \vec{k}\cdot\vec{\Delta r})}\; d\vec{\Delta r}\; d\Delta t & (5.12)\\ &= P(\vec{k}, \omega_D). & (5.13) \end{aligned}$$

This equation defines the power spectrum $P(\vec{k}, \omega_D)$ of a baseband signal $g(\vec{r}, t)$ and is the result of a direct application of the shift property of the Fourier transform.

As an example, let us compute the power spectrum of a plane wave with complex amplitude α, angular frequency ω_0 and with wavevector $\vec{k}_0$. From (5.1) we have

$$g'(\vec{r}, t) = \alpha e^{j(\omega_0 t - \vec{k}_0 \cdot \vec{r})}, \tag{5.14}$$

hence, from (5.6),

$$r'(\vec{\Delta r}, \Delta t) = \sigma_\alpha^2 e^{j(\omega_0 \Delta t - \vec{k}_0 \cdot \vec{\Delta r})} \tag{5.15}$$

where $\sigma_\alpha^2 = E\{|\alpha|^2\}$ is the variance of the amplitude. Finally

$$\begin{aligned} P'(\vec{k}, \omega) &= \iiiint_{-\infty}^{+\infty} \sigma_\alpha^2 e^{j(\omega_0 \Delta t - \vec{k}_0 \cdot \vec{\Delta r})} e^{-j(\omega \Delta t - \vec{k} \cdot \vec{\Delta r})} \, d\vec{\Delta r} \, d\Delta t & (5.16) \\ &= (2\pi)^4 \sigma_\alpha^2 \delta(\omega - \omega_0) \delta(\vec{k} - \vec{k}_0). & (5.17) \end{aligned}$$

Hence, the power spectrum of a plane wave is non-zero at the frequency ω_0 and spatial orientation $\vec{k}_0$ corresponding to the plane wave. This result immediately generalizes to an addition of several plane waves.

5.2.4 Reduced frequencies

Let us consider a baseband signal

$$g(\vec{r}, t) = \alpha e^{j(\omega_D t - \vec{k} \cdot \vec{r})}. \tag{5.18}$$

The angular frequency $\omega_D = 2\pi f_D$ appearing in this expression can be normalized by defining a reduced frequency [127, 145] as in (3.14)

$$\nu_D = \frac{f_D}{f_p} \tag{5.19}$$

where f_p is a sampling frequency, for example the pulse repetition frequency in the case of a pulse-Doppler radar, and a reduced time can be defined as

$$\tau = \frac{t}{T_p} \tag{5.20}$$

where $T_p = \frac{1}{f_p}$ is the sampling period.

Similarly, by defining

$$\vec{\nu}_s = \frac{\vec{u}}{2} = -\frac{\lambda_c}{\pi} \vec{k} \tag{5.21}$$

where λ_c is the wavelength at the carrier frequency. Defining the reduced position vector

$$\vec{\rho} = \frac{2}{\lambda_c} \vec{r} \tag{5.22}$$

allows rewriting (5.18) as

$$g(\vec{\rho}, \tau) = \alpha e^{j2\pi(\nu_D \tau + \vec{\nu}_s \cdot \vec{\rho})}. \tag{5.23}$$

This equation clearly shows that the vector $\vec{\nu}_s = (\nu_{s,x}, \nu_{s,y}, \nu_{s,z})$ defines three reduced spatial frequencies, $\nu_{s,x}$, $\nu_{s,y}$, and $\nu_{s,z}$. In this way, the well-known notion of 2D spatial frequency, e.g., in image processing [151] is extended to a 3D space. As these spatial frequencies denote a direction, $\|\vec{\nu}_s\|$ is always equal to $\frac{1}{2}$. Using these reduced frequencies, the relations between the power spectrum and the covariance function become

$$P(\vec{\nu}_s, \nu_D) = \iiiint_{-\infty}^{+\infty} r(\vec{\Delta\rho}, \Delta\tau) e^{-j2\pi(\nu_D \Delta\tau + \vec{\nu}_s \cdot \vec{\Delta\rho})} \, d\vec{\Delta\rho} \, d\Delta\tau \tag{5.24}$$

and

$$r(\vec{\Delta\rho}, \Delta\tau) = \iiiint_{-\infty}^{+\infty} P(\vec{\nu}_s, \nu_D) e^{j2\pi(\nu_D \Delta\tau + \vec{\nu}_s \cdot \vec{\Delta\rho})} \, d\vec{\nu}_s \, d\nu_D. \tag{5.25}$$

5.2.5 Projection principle

The covariance function[1]

$$r(\vec{\Delta\rho}, \Delta\tau) = r(\Delta\rho_x, \Delta\rho_y, \Delta\rho_z, \Delta\tau) \tag{5.26}$$

corresponds to all possible spatial vector lags $\Delta\rho_x$, $\Delta\rho_y$, and $\Delta\rho_z$ in the 3D space (ρ_x, ρ_y, ρ_z). To make the measurements necessary to compute r at any possible spatial vector lag, a 3D antenna is necessary.

Let us now consider that only a 1D antenna, without loss of generality aligned along the x-axis, is available, permitting only the measurements of all lags along the x axis, i.e., the lags $\vec{\Delta\rho} = (\Delta\rho_x, 0, 0)$. In this case, only the 2D covariance function

$$r_x(\Delta\rho_x, \Delta\tau) = r(\Delta\rho_x, 0, 0, \Delta\tau) \tag{5.27}$$

can be obtained. The power spectral density of this covariance function is given by its 2D Fourier transform

$$P_x(\nu_{s,x}, \nu_D) = \iint_{-\infty}^{+\infty} r_x(\Delta\rho_x, \Delta\tau) e^{-j2\pi(\nu_D \Delta\tau + \nu_{s,x} \Delta\rho_x)} d\Delta\rho_x \, d\Delta\tau. \tag{5.28}$$

and its inverse is given by

$$r_x(\Delta\rho_x, \Delta\tau) = \iint_{-\infty}^{+\infty} P_x(\nu_{s,x}, \nu_D) e^{j(\nu_D \Delta\tau + \nu_{s,x} \Delta\rho_x)} d\nu_{s,x} \, d\nu_D. \tag{5.29}$$

Rewriting (5.25) at $\Delta\rho_y = \Delta\rho_z = 0$ yields

$$r(\Delta\rho_x, 0, 0, \Delta\tau) = \iint_{-\infty}^{+\infty} \left[\iint_{-\infty}^{+\infty} P(\vec{\nu}_s, \nu_D) d\nu_{s,y} \, d\nu_{s,z} \right] e^{j(\nu_D \Delta\tau + \nu_{s,x} \Delta\rho_x)} d\nu_{s,x} \, d\nu_D. \tag{5.30}$$

[1]The baseband function is considered here.

Equating (5.29) and (5.30) yields

$$P_x(\nu_{s,x}, \nu_D) = \iint_{-\infty}^{+\infty} P(\nu_{s,x}, \nu_{s,y}, \nu_{s,z}, \nu_D) d\nu_{s,y}\, d\nu_{s,z} \tag{5.31}$$

which shows that the 2D power spectrum $P_x(\nu_{s,x}, \nu_D)$ is the integral of $P(\vec{\nu}_s, \nu_D)$ along the planes parallel to the plane $\nu_{s,y} = \nu_{s,z} = 0$ and going through each $(\nu_{s,x}, \nu_D)$, as stated in [127, 145]. In other words, $P_x(\nu_{s,x}, \nu_D)$ is the projection of $P(\nu_{s,x}, \nu_{s,y}, \nu_{s,z}, \nu_D)$ on the planes parallel to the plane $\nu_{s,y} = \nu_{s,z} = 0$.

Although the proof is made here for the particular orientation $\vec{1}_x$, its extension to arbitrary orientations is trivial. This property is a variant of the projection-slice theorem [151, Section 8.2], which states that [52] "the Fourier transform of a projection is a slice of the 2D Fourier transform from the region from which the projection was obtained" where $P_x(\nu_{s,x}, \nu_D)$ is the projection of $P(\vec{\nu}_s, \nu_D)$ onto the $(\nu_{s,x}, \nu_D)$ plane and $r_x(\Delta\rho_x, \Delta\tau)$ is the slice of $r(\vec{\Delta\rho}, \Delta\tau)$ at $(\Delta\rho_y, \Delta\rho_z) = (0, 0)$.

5.3 Clutter power spectrum locus

Now, let us consider the signal components due to the clutter.

5.3.1 4D Clutter power spectrum locus

The clutter is modeled as the superposition of a large number of independent clutter sources [60, 173] located along the isorange of interest as discussed in Section 3.5.2. Furthermore, let us assume that no range ambiguity occurs. Each clutter patch contributes with a signal corresponding to a distinct direction of arrival $\vec{\nu}_s$. Hence, the signal from each clutter patch corresponds to a distinct point in the spatio-temporal frequency domain $(\vec{\nu}_s, \nu_D)$. This can be thought of as if the isorange in the 3D spatial domain was imaged into another curve in the 4D frequency domain $(\vec{\nu}_s, \nu_d)$. For this reason, let us call this 4-dimensional curve in the frequency domain the *4D clutter power spectrum locus* [127, 145]. Note that this curve is independent of the characteristics of the antenna. Figures 5.2 and 5.3 illustrate the 4D clutter power spectrum locus. The representation consists of two graphs. The first graph is a projection in the 3D space $(\nu_{s,x}, \nu_{s,y}, \nu_D)$, while the second graph is a projection in the 3D space $(\nu_{s,x}, \nu_{s,y}, \nu_{s,z})$ of the spatial frequencies. As the norm of $\vec{\nu}_s$ is constant and equal to $\frac{1}{2}$, the latter representation of the projection of the 4D clutter power spectrum locus yields a curve on a (3D) sphere with a radius equal to $\frac{1}{2}$.

5.3.2 2D Clutter power spectrum locus

In the case of a linear antenna, it is well known that the clutter power spectrum exhibits a so-called clutter ridge and that, in the limit for an infinitely long continuous antenna and for an infinitely long continuous observation interval, the 2D power spectrum is concentrated along a 2D curve. This curve which can also be obtained by physical reasoning, is called angle-Doppler curve,

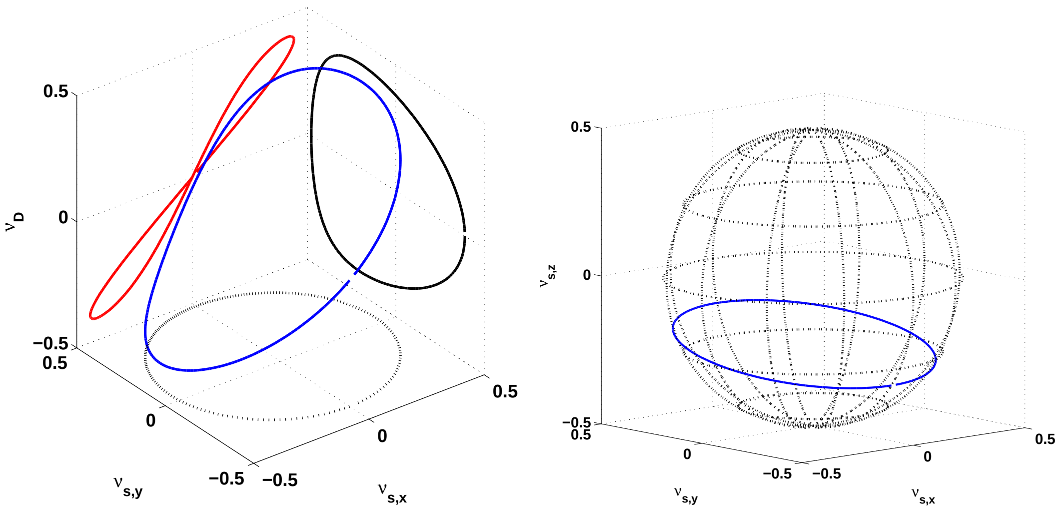


Figure 5.2: 4D clutter power spectrum locus (blue curve) (wing-to-wing bistatic formation).

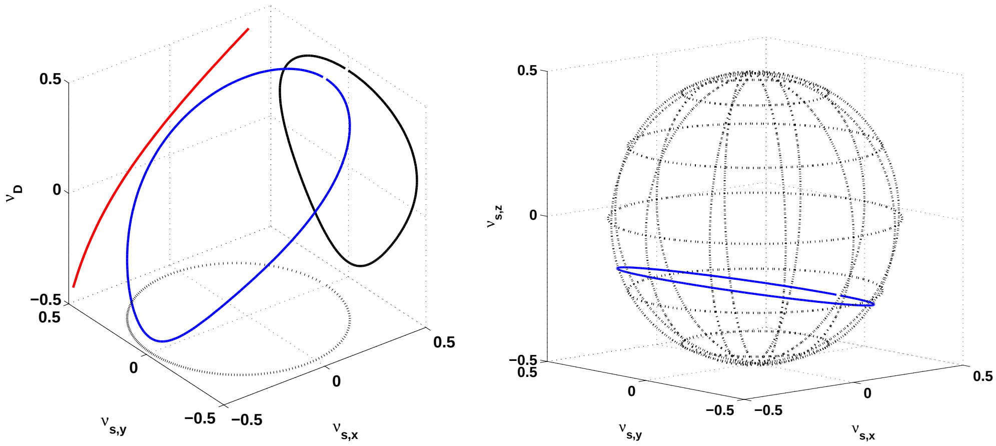
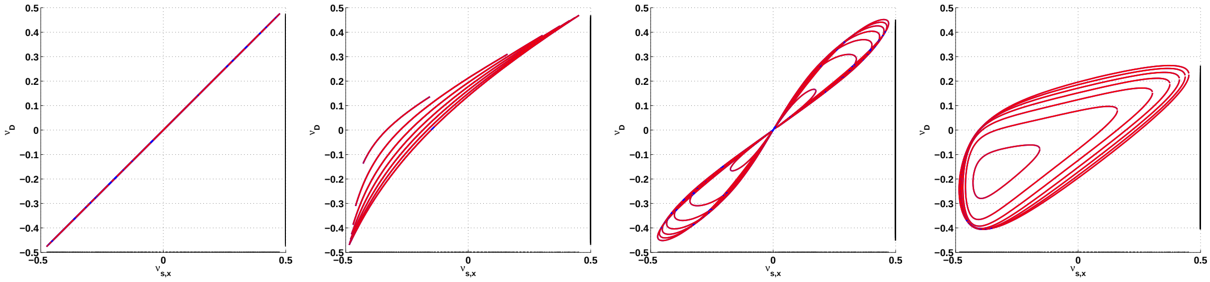


Figure 5.3: 4D clutter power spectrum locus (blue curve) (in trail bistatic formation).

direction-Doppler (DD) curve, or (2D) clutter power spectrum locus [70, 95, 124, 145, 173, 178]. This curve is also called the clutter ridge by language abuse. The shape of this 2D clutter power spectrum locus varies in a complex way with changes in the geometric configuration [95, 108, 178]. Examples are shown in Figure 5.4.

The function $P_x(\nu_{s,x}, \nu_D)$ being the projection of $P(\vec{\nu}_s, \nu_D)$, it immediately follows that the 2D clutter power spectrum locus (a 2D curve) is the projection of the 4D clutter power spectrum locus (a 4D curve). In other words, once the 4D clutter power spectrum locus in the $(\vec{\nu}_s, \nu_D)$ space is known, its 2D counterpart in any 2D plane can be obtained immediately, e.g., that corresponding to $(\nu_{s,x}, \nu_D)$, which is the customary 2D clutter power spectrum locus. Therefore, many of the complex behaviors can now be understood in terms of the projection of the 4D clutter power spectrum locus on a 2D plane. For example, the effect of a non-zero crab angle (angle



(a) (b) (c) (d)

Figure 5.4: 2D clutter power spectrum loci at different ranges (a) for a monostatic radar configuration and for three bistatic configurations: (b) for aircrafts in trail formation (c) for aircrafts in wing-to-wing formation and (d) for aircrafts flying on orthogonal trajectories.

between $\nu_{s,x}$ and the velocity vector assumed horizontal) on the 2D clutter power spectrum locus is difficult to interpret, although it simply results in a rotation of the 4D clutter power spectrum

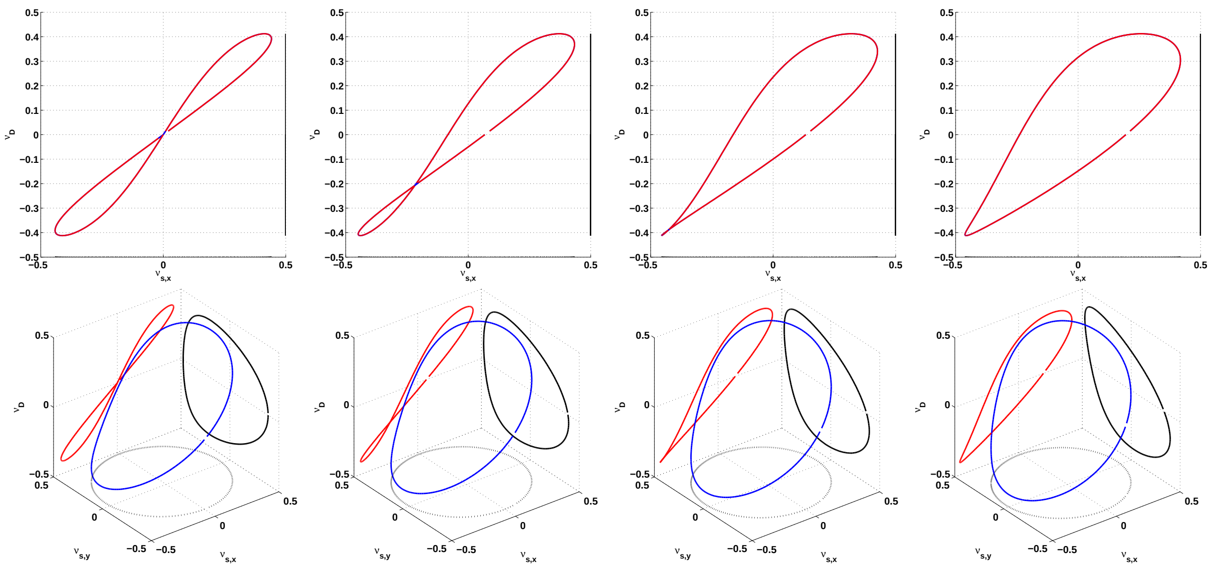


Figure 5.5: Clutter power spectrum loci for a wing-to-wing bistatic configuration for crab angles of, from left to right, 0^0, 10^0, 20^0 and 30^0.

locus around an axis parallel to the k_z axis as illustrated in Figure 5.5 where the top row depicts the 2D clutter power spectrum locus and the bottom row depicts one 3D projection of the 4D clutter power spectrum locus for increasing crab angles.

5.4 Clutter covariance matrix spectrum

5.4.1 Introduction

In real radar systems, only samples $g(\vec{\rho}_n, \tau_m)$ of the signal field $g(\vec{\rho}, \tau)$ are available. The temporal samples τ_m correspond to the time at which the pulses[2] are transmitted. For a constant pulse repetition interval (PRI) T_p, $\tau_m = m$. The spatial sampling corresponds to the actual location of the array elements. For a uniform linear array (ULA) aligned with the x-axis, $\vec{\rho}_n = n\ \frac{2}{\lambda}d\ \vec{1}_x$ where d is the distance between two adjacent antenna elements and $\vec{1}_x$ is a unit vector aligned on the x-axis. The set of $N \times M$ samples of a particular range gate, measured at the N antenna elements resulting from the M pulses is called a snapshot and is usually lexicographically-ordered in a vector $\mathbf{y}$ with

$$\mathbf{y}(n + Nm) = y(\vec{\rho}_n, \tau_m) \tag{5.32}$$

where $\vec{\rho}_n$ denotes the (reduced) position of the n^{th} sensor and τ_m the (reduced) time of the m^{th} temporal sample.

5.4.2 Relationship between the covariance matrix and the covariance function

In all generality, the covariance matrix of a zero-mean data vector $\mathbf{y}$ is given by

$$\mathbf{R} = E\{\mathbf{y}\mathbf{y}^\dagger\} \tag{5.33}$$

or, equivalently, one element of this matrix is given by

$$R(n + Nm, n' + Nm') = E\{y(\vec{\rho}_n, \tau_m)y^*(\vec{\rho}_{n'}, \tau_{m'})\}. \tag{5.34}$$

The definition (5.10) of the covariance function immediately yields, for a wide-sense stationary signal $y(\vec{\rho}, \tau)$

$$R(n + Nm, n' + Nm') = r(\vec{\rho}_n - \vec{\rho}_{n'}, \tau_m - \tau_{m'}). \tag{5.35}$$

In all generality, aside from being hermitian, the covariance matrix does not exhibit a particular structure. If a constant PRF equal to f_p is used, $\tau_m - \tau_{m'} = m - m'$ and the covariance matrix exhibits a block-Toeplitz structure. If a ULA is used, $\vec{\rho}_n - \vec{\rho}_{n'} = \frac{2}{\lambda}d\ (n - n')\ \vec{1}_x$ where $\vec{1}_x$ is a unit vector aligned on the antenna axis. In the latter case, if a constant PRF is used, the covariance matrix exhibits a Toeplitz-block-Toeplitz structure.

5.4.3 Estimation of the clutter power spectrum

The power spectrum of the signal y can be estimated from the covariance matrix of $\mathbf{y}$ by using, e.g., the discrete time Fourier transform or the minimum variance estimator (MVE) [26, 171]. If the clairvoyant covariance matrix (5.33) is used, the estimation error on the power spectrum

[2] The notion of "pulses" is generalized in Chapter 6.

is only due to the fact that a finite number of samples is available to perform the power spectrum estimation. Additional estimation errors occur on the power spectrum if an estimate of the covariance matrix is used instead of the clairvoyant one. To simplify the discussion, and unless otherwise noted, the clairvoyant covariance matrix is used in the remainder of this chapter.

5.4.3.1 Fourier-transform based spectral estimation

The signal-match estimator [84]

$$P_{\mathrm{SM}}(\vec{\nu}_s, \nu_D) = \frac{\mathbf{v}^\dagger(\vec{\nu}_s, \nu_D)\mathbf{R}\mathbf{v}(\vec{\nu}_s, \nu_D)}{\mathbf{v}^\dagger(\vec{\nu}_s, \nu_D)\mathbf{v}(\vec{\nu}_s, \nu_D)} \tag{5.36}$$

is motivated by the fact that the steering vector given by

$$\begin{aligned}\mathbf{v}(\vec{\nu}_s, \nu_D) = [&e^{j2\pi(\nu_D\tau_0+\vec{\nu}_s\cdot\vec{\rho}_0)}, e^{j2\pi(\nu_D\tau_0+\vec{\nu}_s\cdot\vec{\rho}_1)}, \ldots, e^{j2\pi(\nu_D\tau_0+\vec{\nu}_s\cdot\vec{\rho}_{N-1})}, \\ &e^{j2\pi(\nu_D\tau_1+\vec{\nu}_s\cdot\vec{\rho}_0)}, \ldots, e^{j2\pi(\nu_D\tau_{M-1}+\vec{\nu}_s\cdot\vec{\rho}_{N-1})}]^T\end{aligned} \tag{5.37}$$

is "matched" to the signal — the plane wave — the power of which must be estimated. This power P_{SM} is related to the periodogram [80, 105]

$$P_{\mathrm{PER}}(\vec{\nu}_s, \nu_D) = \frac{1}{NM}|\mathbf{v}^\dagger(\vec{\nu}_s, \nu_D)\mathbf{y}|^2 \tag{5.38}$$

as $P_{\mathrm{SM}} = E\{P_{\mathrm{PER}}\}$. Note that the periodogram is a random quantity while P_{SM} is deterministic.

Now, let us relate the estimate P_{SM} to the power spectrum of the continuous signal y. (5.36) yields

$$P_{\mathrm{SM}}(\vec{\nu}_s, \nu_D) = \frac{1}{NM}\sum_{n=0}^{N-1}\sum_{m=0}^{M-1}\sum_{n'=0}^{N-1}\sum_{m'=0}^{M-1} e^{-j2\pi[\nu_D(\tau_m-\tau_{m'})+\vec{\nu}_s\cdot(\vec{\rho}_n-\vec{\rho}_{n'})]}R(n+Nm, n'+Nm'). \tag{5.39}$$

Taking into account (5.35), the inverse Fourier transform of the power spectrum (5.25), and using $\vec{\Delta\rho} = \vec{\rho}_n - \vec{\rho}_{n'}$ and $\Delta\tau = \tau_m - \tau_{m'}$, (5.39) can be rewritten as

$$\begin{aligned}P_{\mathrm{SM}}(\vec{\nu}_s, \nu_D) = \frac{1}{NM}&\iiiint_{-\infty}^{+\infty} P(\vec{\nu}'_s, \nu'_D) \\ &\sum_{n=0}^{N-1}\sum_{m=0}^{M-1}\sum_{n'=0}^{N-1}\sum_{m'=0}^{M-1} e^{j2\pi[\nu'_D(\tau_m-\tau_{m'})+\vec{\nu}'_s\cdot(\vec{\rho}_n-\vec{\rho}_{n'})]}e^{-j2\pi[\nu_D(\tau_m-\tau_{m'})+\vec{\nu}_s\cdot(\vec{\rho}_n-\vec{\rho}_{n'})]}\ d\vec{\nu}'_s\ d\nu'_D.\end{aligned} \tag{5.40}$$

The sum of the exponential factors of the integrand can be rewritten by using the spatio-temporal beampattern [26, 171]

$$B(\vec{\nu}_s, \nu_D, \vec{\nu}'_s, \nu'_D) = \frac{1}{\sqrt{NM}}\sum_{n=0}^{N-1}\sum_{m=0}^{M-1} e^{j2\pi[(\nu_D-\nu'_D)\tau_m+(\vec{\nu}_s-\vec{\nu}'_s)\cdot\vec{\rho}_n]} \tag{5.41}$$

$$= \frac{1}{\sqrt{NM}}\mathbf{v}^\dagger(\vec{\nu}'_s, \nu'_D)\mathbf{v}(\vec{\nu}_s, \nu_D) \tag{5.42}$$

that can be interpreted as the response of the receive array to a plane wave with spatio-temporal frequencies $(\vec{\nu}_s, \nu_D)$ as the receive array is steered in a direction defined by the spatio-temporal frequencies $(\vec{\nu}'_s, \nu'_D)$. It is important to note that the beampattern only depends on the difference between these spatio-temporal frequencies, that is

$$B(\vec{\nu}_s, \nu_D, \vec{\nu}'_s, \nu'_D) = B_0(\vec{\nu}_s - \vec{\nu}'_s, \nu_D - \nu'_D) = B(\vec{\nu}_s - \vec{\nu}'_s, \nu_D - \nu'_D, 0, 0). \tag{5.43}$$

Hence (5.40) becomes

$$P_{\mathrm{SM}}(\vec{\nu}_s, \nu_D) = \iiiint_{-\infty}^{+\infty} P(\vec{\nu}'_s, \nu'_D) |B(\vec{\nu}_s, \nu_D, \vec{\nu}'_s, \nu'_D)|^2 d\vec{\nu}'_s \, d\nu'_D \tag{5.44}$$

or yet

$$P_{\mathrm{SM}}(\vec{\nu}_s, \nu_D) = \iiiint_{-\infty}^{+\infty} P(\vec{\nu}'_s, \nu'_D) |B_0(\vec{\nu}_s - \vec{\nu}'_s, \nu_D - \nu'_D)|^2 d\vec{\nu}'_s \, d\nu'_D. \tag{5.45}$$

This last expression shows that the estimate of the power spectrum obtained using the signal-match operator is equal to the actual power spectrum $P(\vec{\nu}_s, \nu_D)$ convolved with $|B_0|^2$.

5.4.3.2 Minimum variance spectral estimation

The minimum variance estimator [80, 84] is given by

$$P_{\mathrm{MVE}}(\vec{\nu}_s, \nu_D) = \frac{MN}{\mathbf{v}^\dagger(\vec{\nu}_s, \nu_D)\mathbf{R}^{-1}\mathbf{v}(\vec{\nu}_s, \nu_D)} \tag{5.46}$$

and can be rewritten as

$$P_{\mathrm{MVE}}(\vec{\nu}_s, \nu_D) = \mathbf{c}^\dagger(\vec{\nu}_s, \nu_D)\mathbf{R}\mathbf{c}(\vec{\nu}_s, \nu_D) \tag{5.47}$$

with

$$\mathbf{c}(\vec{\nu}_s, \nu_D) = \sqrt{MN}\frac{\mathbf{R}^{-1}\mathbf{v}(\vec{\nu}_s, \nu_D)}{\mathbf{v}^\dagger(\vec{\nu}_s, \nu_D)\mathbf{R}^{-1}\mathbf{v}(\vec{\nu}_s, \nu_D)}. \tag{5.48}$$

The filter $\mathbf{c}$ is recognized as the optimum filter, minimizing the output power while keeping the power of a plane wave with spatio-temporal frequencies $(\vec{\nu}_s, \nu_D)$ unaffected. The factor MN in (5.46) and $\sqrt{MN}$ in (5.48) are needed for P_{MVE} to reflect the power spectral density [80].

As in the previous section, defining a MVE beampattern yields

$$B_{\mathrm{MVE}}(\vec{\nu}_s, \nu_D, \vec{\nu}'_s, \nu'_D) = \mathbf{c}(\vec{\nu}_s, \nu_D)^\dagger \mathbf{v}(\vec{\nu}'_s, \nu'_D) \tag{5.49}$$

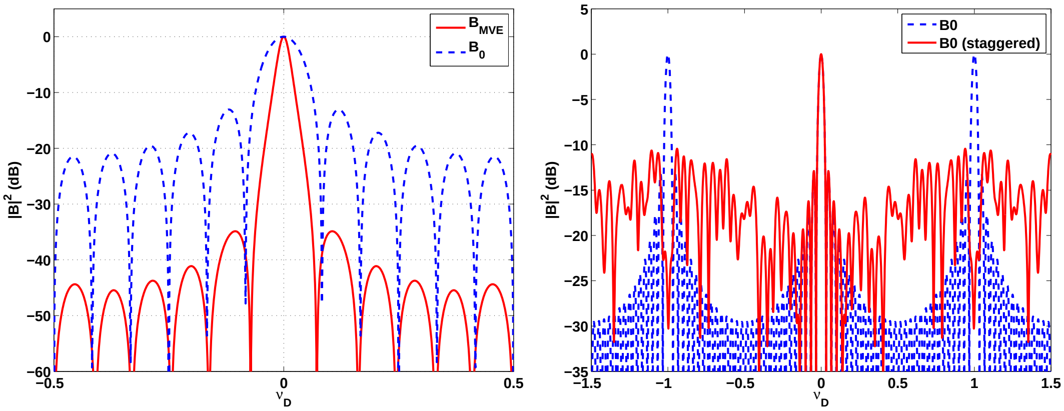
and expressing the relation between P_{MVE} and the power spectrum P as a superposition integral yields

$$P_{\mathrm{MVE}}(\vec{\nu}_s, \nu_D) = \iiiint_{-\infty}^{+\infty} P(\vec{\nu}'_s, \nu'_D) |B_{\mathrm{MVE}}(\vec{\nu}_s, \nu_D, \vec{\nu}'_s, \nu'_D)|^2 \, d\vec{\nu}'_s \, d\nu'_D. \tag{5.50}$$

This expression again shows that the estimate of the power spectrum obtained from the covariance matrix using the MVE is related to the power spectrum of the signal through a superposition integral with a space-time varying kernel.

5.4.4 Effect of the sampling

The beampatterns B and B_{MVE} for the two estimators analyzed in the previous section are maximum for $(\vec{\nu}_s, \nu_D) = (\vec{\nu}_s', \nu_D')$. This is illustrated in Figure 5.6 (a), where the evolutions of $|B(\vec{\nu}_s, \nu_D, \vec{\nu}_s', \nu_D')|^2$ and $|B_{\mathrm{MVE}}(\vec{\nu}_s, \nu_D, \vec{\nu}_s', \nu_D')|^2$ respectively are given for $\vec{\nu}_s = \vec{\nu}_s'$ and $\nu_D' = 0$. $(\vec{\nu}_s', \nu_D')$ is selected such that $P(\vec{\nu}_s', \nu_D') \neq 0$ in order to illustrate the high resolution property of the MVE. Although $P(\vec{\nu}_s, \nu_D)$ does not depend on the particular receive array layout and



(a) (b)

Figure 5.6: Normalized (dividing B and B_{MVE} by $\sqrt{MN}$) beampattern ($\nu_D' = 0$, $\vec{\nu}_s = \vec{\nu}_s'$) (a) for the two spectral estimators described in the text (M=12) and (b) in the case of a non-uniform temporal sampling (M=40).

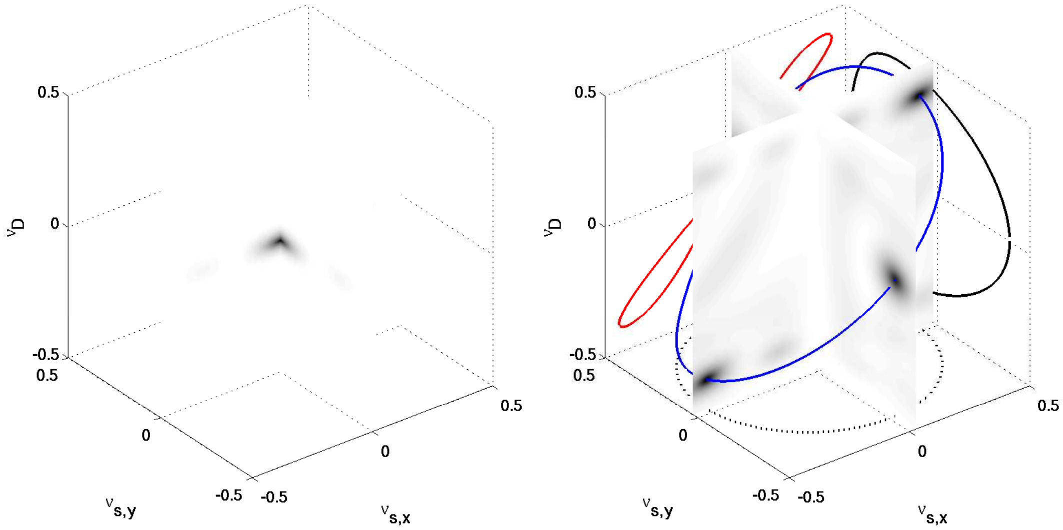
on the considered temporal sampling, the beampattern and hence the estimated power spectra $P_{\mathrm{MVE}}(\vec{\nu}_s, \nu_D)$ and $P_{\mathrm{SM}}(\vec{\nu}_s, \nu_D)$ do.

If a uniform sampling is considered, the beampattern exhibits temporal grating lobes, i.e., aliasing, as illustrated by the curve B_0 in Figure 5.6 (b). As it is well known [106, 171], these grating lobes essentially disappear, at the expense of higher sidelobes, if non-uniform sampling, i.e. pulse staggering, is considered as illustrated in Figure 5.6 (b).

5.4.5 Discussion

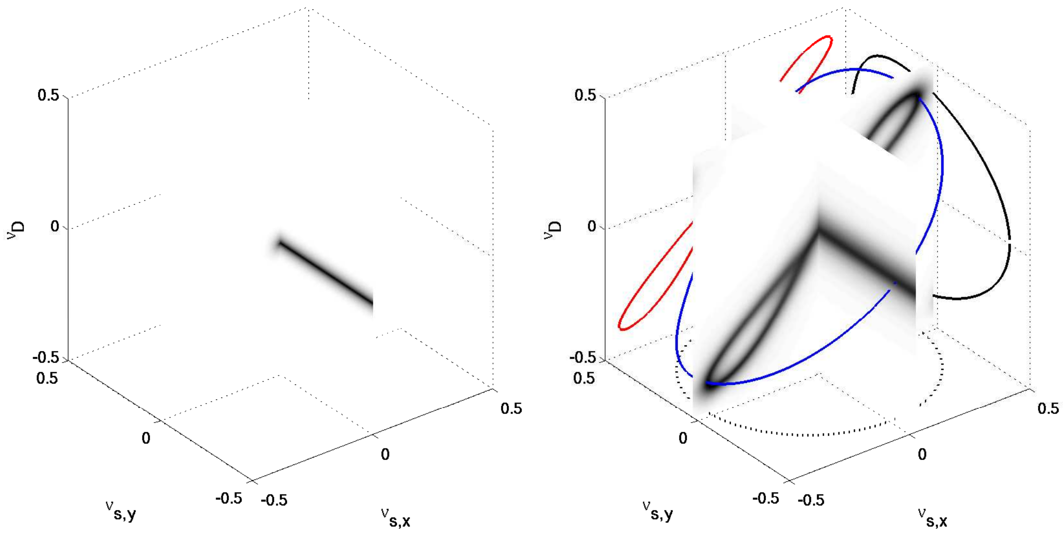
The relation between the power spectrum of the continuous signal $y(\vec{\rho}, \tau)$ and the power spectrum estimated from the covariance matrix $\mathbf{R}$ provides with the formal link between the clutter power spectrum locus described in the previous section and the estimate of the clutter power spectrum obtained from sampled data. This is illustrated in Figure 5.7 where the kernel corresponding to the MVE in the case of a 12 elements circular antenna is shown together with the corresponding clutter power spectrum. The continuous functions of $(\vec{\nu}_s, \nu_D)$ are represented by two cuts along the planes $\nu_{s,y} = 0$ and $\nu_{s,x} = 0$ respectively.

Obviously, the clutter power spectrum is concentrated along the clutter power spectrum locus. Similarly, Figure 5.8 shows the kernel and the clutter power spectrum in the case of a ULA. As



(a) (b)

Figure 5.7: (a) Kernel (in gray scale) linking the MVE power spectrum based on $\mathbf{R}$ and the power spectrum of y. (b) Comparison between the estimated clutter power spectrum and the clutter power spectrum locus (blue line) for a 12-element circular antenna.



(a) (b)

Figure 5.8: (a) Kernel (in gray scale) linking the MVE power spectrum based on $\mathbf{R}$ and the power spectrum of y. (b) Comparison between the estimated clutter power spectrum and the clutter power spectrum locus (blue line) for a 12-element ULA.

expected, in this case, the clutter power spectrum does not depend on the spatial frequencies components along the y and z axis. It only varies with ν_x and ν_D.

5.5 Range-dependence of the clutter power spectrum

Range-dependence of the clutter covariance matrix is a major issue in STAP as it makes its estimation more complex. Therefore, let us analyze the range-dependence of the clutter covariance matrix in the spectral domain. The link between the clutter power spectrum locus and the clutter power spectrum was analyzed in the previous section which permits to study the behavior of the clutter power spectrum locus (a curve) as a function of the configuration parameters (the range for instance) and from this study deduce the behavior of the clutter power spectrum. This problem is simpler than a direct analysis of the clutter power spectrum. A necessary condition of range independence of the clutter power spectrum is thus that the clutter power spectrum locus is range independent.

For monostatic scenarios and for most bistatic scenarios, the 4D clutter power spectrum locus depends on the considered range. Figure 5.9 illustrates this fact for a bistatic scenario and Figure 5.10 for a monostatic scenario. For monostatic scenarios, the 4D clutter power spectrum

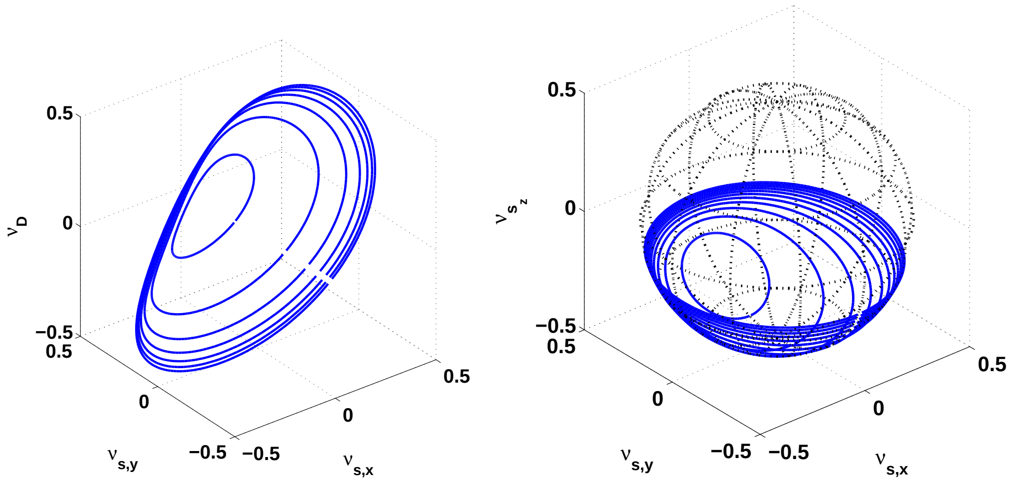


Figure 5.9: Evolution of the 4D clutter power spectrum locus for increasing ranges in the case of a wing-to-wing bistatic scenario.

locus is located on a 3D hyperplane. The consequence is that the projection of the clutter power spectrum locus on a suitably oriented 2D plane yields overlapping straight lines at any range. This happens if the considered 2D plane is oriented parallel to the platform velocity vector. Indeed, this means that the linear array must be parallel to the velocity vector if range dependence must be avoided.

The conditions under which the 2D power spectrum locus is independent of the range, as derived in [92] in the case of a ULA, are a constant PRF and a horizontal velocity. This derivation is based on an explicit expression of the equation of the 2D clutter power spectrum locus. By

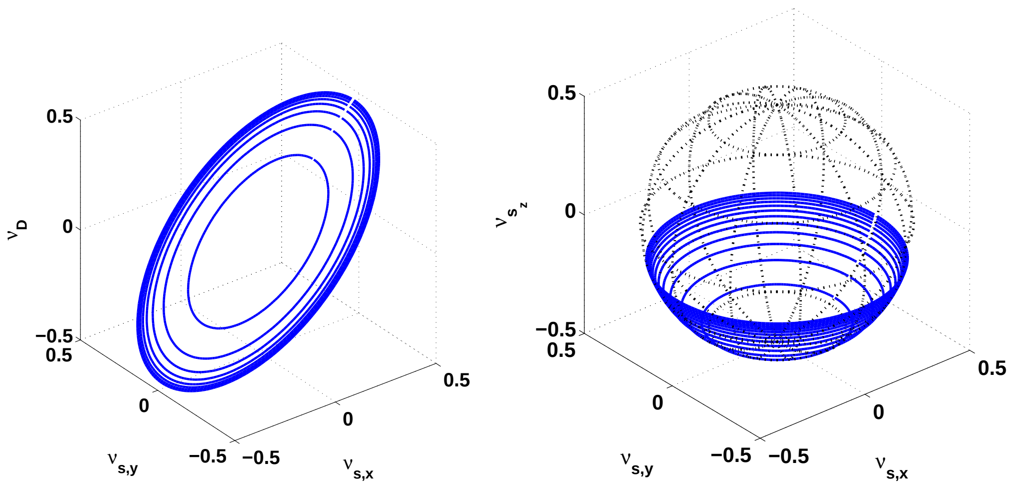


Figure 5.10: Evolution of the 4D clutter power spectrum locus for increasing ranges in the case of a monostatic scenario.

using the 4D clutter power spectrum locus, the same conclusions can be obtained by a simple argument. First, let us consider the spatial frequencies. In order to obtain a range-independent clutter power spectrum locus, the curves at different ranges need to overlap in the 3D space of the spatial frequencies. As the elevation of points along the curves in the 3D space of the spatial frequencies only depends on the elevation angle at which the scatterers along the isorange are seen from the receiver, an overlap in the 3D space of the spatial frequencies occurs if and only if the receiver is on the (flat) ground. In this case, the scatterers are seen at an elevation angle equal to zero, regardless of the range. Second, let us consider the Doppler frequency. Let us require the range-independence of the Doppler frequency corresponding to a particular spatial frequency. In a configuration for which the receiver is located on the ground, the Doppler frequency shift due to the receiver velocity is constant along radial lines from the receiver. The only configurations for which the Doppler frequency shift induced by the transmitter velocity is independent of range are either

- when the transmitter is static (including no vertical velocity component), in which case this Doppler frequency is zero and hence range-independent, or
- when the transmitter is located on the ground and at the same location as the receiver, in which case the Doppler frequency only depends on the transmitter azimuth angle (and on the transmitter velocity) which is range-independent. Note that, in this case, the velocity of the transmitter may be different from that of the receiver.

In the case of a range-dependence of the 4D clutter power spectrum locus, an estimate of the clutter covariance matrix by using the sample covariance matrix is biased as the averaged snapshots are not identically distributed. This is illustrated in the case of a 12-elements circular antenna in Figure 5.11, where the clutter power spectrum estimated from the sample covariance matrix is presented. As can be seen, the power spectrum significantly deviates from the true clutter power

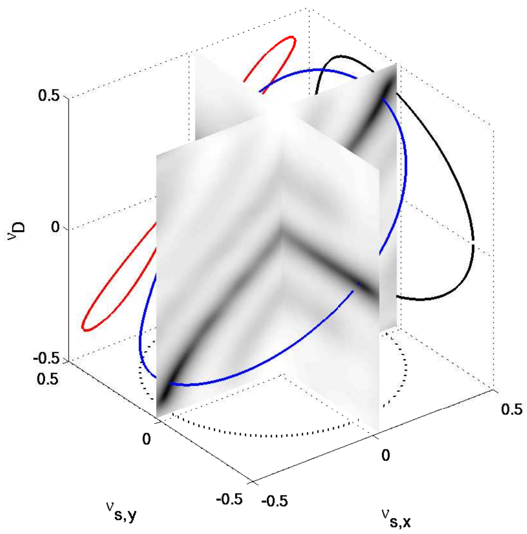


Figure 5.11: MVE of the power spectrum estimated from the sample covariance matrix.

spectrum estimate based on the clairvoyant covariance matrix depicted in Figure 5.7. In particular, the estimated power spectrum is not concentrated along the clutter power spectrum locus. Ways to overcome this problem are presented in Chapter 7.

5.6 Extension to the multistatic case

If the clutter from different bistatic radars is assumed independent, the clutter covariance matrix is block-diagonal as discussed in Section 3.5.2.4. This covariance matrix is thus *not* composed of samples of a covariance function. Hence, the whole concept of the power spectrum of a covariance matrix is lost. Regardless of this, if the different platforms composing the multistatic radar exhibit different velocities, a single Doppler frequency cannot be defined. Hence decomposing the signal in a basis of (sampled) exponential does not yield any meaningful result as noted in [128]. The fact that, for a bistatic radar, the steering vectors are indeed Kronecker products of sampled exponentials and, in particular for a ULA and with uniform temporal sampling, the steering vectors are equal to the basis vectors of the Fourier transform explains the usefulness of the Fourier transform in analyzing the behavior of the clutter covariance matrix.

Note that the obvious solution consists in decomposing the signal in a basis of steering vectors considering the target velocity vector as the independent quantity, as opposed to the use of the Doppler frequency as the independent quantity. However, doing so also hides the Doppler-frequency range-dependence of the clutter covariance matrix making it impossible to analyze this range-dependence.

Chapter 6

Transmitters of opportunity as signal source

6.1 Introduction

This chapter considers the feasibility of using transmitters of opportunity to perform moving target detection by using STAP.

Usually, the waveform transmitted by radars used to perform moving target detection is a coherent pulse train. The achievable detection range is linked to the amount of transmitted energy. Moreover, the range resolution is linked to the pulse length, i.e., the time during which the energy is transmitted. Hence achieving simultaneously a high energy and a high range resolution with a coherent pulse train is technically challenging as it requires very large peak powers. Pulse compression [133, 146] can be used in order to cope with power and resolution issues. In this case, instead of transmitting a pulse train, a modulated signal is transmitted, in order to spread the transmitted power over a duration which is much longer than the pulse length. The received echoes are then processed to "compress" the transmitted signal and to achieve the desired range resolution. This processing consists in a matched filtering, already described in Chapter 4.

More recently, and due to the availability of fast digital-to-analog converters, some developments in the field of waveform optimization have been conducted [17, 139]. The idea of waveform optimization consists in optimizing the transmitted waveform in order to minimize the effect of the interferences on the detection performance. These methods assume the presence of a dedicated transmitter and full control of the radar waveform by the operator.

Noise radars [7, 8, 9, 16, 118, 119], for which the transmitted waveform is an truly noise-like, have been developed. The main advantage of these waveforms is their covertness, electronic counter-counter measures (ECCM) capabilities and the relative absence of ambiguities [8, 16]. Nevertheless, the sidelobes of the ambiguity function are larger than in the case of a standard coherent pulse train [146], which however can be solved by using the periodic ambiguity function [101, 133], provided that the transmitted noise signal is periodic.

As far as transmitters of opportunity are concerned, the transmitted waveform depends on the transmitter. Some of these transmitters transmit waveforms that have properties close to those of noise. This approach however differs from the approach described in the previous paragraphs as the objective is not as much to design a waveform that maximizes the performance. The intention is rather to determine the performance that can be achieved by using the considered transmitted waveform, which is precisely what the following sections analyze.

Transmitters of opportunity are transmitters present in the environment and not dedicated to radars. Considered transmitters include analog TV [58], FM radio [179], digital terrestrial TV [152], mobile phone (GSM) base stations [87, 162, 163, 164], WiFi beacons [63], satellites on low-earth orbit [31, 57, 74], and geostationary digital TV broadcast satellites [29, 140]. Features of interest for the selection of the transmitter type include spatial and temporal coverage (revisit time), power, central frequency, and bandwidth of the transmitted signal, as well as shape of the ambiguity function. The bandwidth determines the achievable range-resolution. The shape of the ambiguity function is essential for determining the detection performance of the radar. In particular, signals from digital modulation (GSM, DVB) have much less range and Doppler ambiguities than other modulations [56], which makes them more suitable for passive radar.

The remainder of this chapter carries out theoretical developments in the most general case as far as possible. To illustrate these developments, GSM base stations as illuminators of opportunity are considered. They have an ubiquitous spatial coverage, are permanent in time and have a thumbtack-like ambiguity function due to the noise-like behavior of the GMSK modulation. The main drawback of GSM base station signals is the small bandwidth [163] that yields a range resolution of about 1.8 km. Thus, a GSM-based radar can only be used to perform moving target detection. The Doppler frequency resolution only depends on the coherent processing interval (CPI). A CPI of a few tenths of a second is easily achievable and yields a Doppler frequency resolution of a few Hz.

The following developments first consider the single channel case, review the ambiguity function and introduce the generalized ambiguity function, motivated as the response of a filter designed to optimally reject interferences. Finally, the extension of the single channel case to the multichannel case makes spatio-temporal processing, hence STAP, possible.

6.2 Single-channel, clutter-free case

6.2.1 Introduction

This section considers a radar with a single receive channel and develops the expressions for the range and Doppler-frequency resolution in the case of phase-modulated waveforms. The quantities related to a single channel are denoted by the subscript u.

6.2.2 Signal model

Let us consider a transmitter transmitting a signal with a complex envelope $p(t)$, a bandwidth B and modulated at a carrier angular frequency ω_c. A point target is located at range R_T from the transmitter and range R_R from the receiver, has a complex reflectivity α_c and is moving with a velocity such that the total range-rate is v. The complex envelope of the signal echoed by the target and received by the receiver is

$$\alpha_c\, p(t-\tau_a(t))e^{j\omega_{D_a}t}+n(t) \tag{6.1}$$

where $n(t)$ is the thermal noise and

$$\tau_a(t)=\frac{R_T+R_R}{c}+\frac{v}{c}t \tag{6.2}$$

is the time-dependent delay between the transmission of the signal and the reception of its echo. The subscript a denotes that the value τ_a is the *actual* value of the quantity, and $\omega_{D_a}=2\pi f_D$ where f_D is the total Doppler frequency shift induced by the target range rate.

Let us assume that the transmitted signal is not significantly deformed at the reception due to the motion of the target during the CPI, which is the case if the point target remains in the same resolution cell during the whole CPI, i.e., if no range-migration occurs. This condition imposes a limit on the length of the CPI, the range-rate and the bistatic range. During a CPI of duration T_{CPI}, the target bistatic range increases with vT_{CPI} and the corresponding delay is increased with $\frac{v}{c}T_{\text{CPI}}$. If the bandwidth of the transmitted complex envelope $p(t)$ is denoted by B, the condition can be written

$$BT_{\text{CPI}}<\frac{c}{|v|}. \tag{6.3}$$

In this case, the temporal dependence of τ_a in (6.2) can be neglected. Then, the complex envelope of the received signal becomes

$$\tilde{y}(t)=\alpha_c\, p(t-\tau_a)e^{j\omega_{D_a}t}+n(t) \tag{6.4}$$

where the tilde denotes the fact that the complex envelope of the transmitted signal differs from the common pulse train.

If the signal $\tilde{y}(t)$ is sampled at frequency f_s, the i^{th} sample is given by

$$\tilde{y}(i)=\alpha_c\, p(i\Delta t-\tau_a)e^{j\omega_{D_a}i\Delta t}+n(i\Delta t) \tag{6.5}$$

where $\Delta t=1/f_s$ and the sampling frequency f_s is larger than the bandwidth of $p(t)$ such that no aliasing occurs. If M' samples are taken, defining the vectors $\tilde{\mathbf{y}}'_u$, $\mathbf{p}$, $\mathbf{b}'$, and $\mathbf{n}'$ respectively as

$$\tilde{\mathbf{y}}'_u(\tau_a,\nu_{D_a})=[y(0),y(\Delta t),\ldots,y((M'-1)\Delta t)]^T \tag{6.6}$$

$$\mathbf{p}(\tau_a)=[p(-\tau_a),p(\Delta t-\tau_a),\ldots,p((M'-1)\Delta t-\tau_a)]^T \tag{6.7}$$

$$\mathbf{b}'(\nu'_{D_a})=[1,e^{j2\pi\nu'_{D_a}},\ldots,e^{j2\pi\nu'_{D_a}(M'-1)}]^T \tag{6.8}$$

$$\mathbf{n}'=[n(0),n(\Delta t),\ldots,n((M'-1)\Delta t)]^T \tag{6.9}$$

where $\nu'_{D_a}=\frac{f_D}{f_s}$ is the reduced Doppler frequency, yields

$$\tilde{\mathbf{y}}'_u(\tau_a,\nu'_{D_a})=\alpha_c\mathbf{p}(\tau_a)\circ\mathbf{b}'(\nu'_{D_a})+\mathbf{n}'=\tilde{\mathbf{y}}'_{t_u}(\tau_a,\nu'_{D_a})+\mathbf{n}' \tag{6.10}$$

where $\tilde{\mathbf{y}}'_{t_u}$ is the noise-free target signal and the operator $\circ$ denotes the Hadamard product (element-wise product) of its two arguments.

$\mathbf{b}'(\nu'_{D_a})$ has the same shape as the temporal steering vector defined in (3.19). There is however a difference between (3.19) and (6.8) as the former results from a sampling at the pulse repetition frequency (PRF) f_p while the latter results from a sampling at f_s, a frequency at least higher than the bandwidth of $p(t)$ and typically much higher than the pulse repetition frequency. Indeed, the maximum sensible value for the pulse repetition frequency in (3.19) is equal to twice the maximum expected Doppler frequency. Quantities related to the sampling frequency f_s are denoted by an apostrophe, to distinguish them from the quantities related to the usual PRF sampling frequency f_p.

6.2.3 Ambiguity function

The sufficient statistic T needed to determine whether a target is present at the hypothesized (denoted by the subscript h) delay τ_h and Doppler frequency ν'_{D_h} is (see Chapter 4)

$$T(\tau_h, \nu'_{D_h}) = |\tilde{\mathbf{w}}_u^{\dagger}(\tau_h, \nu'_{D_h})\tilde{\mathbf{y}}'_u|^2 \tag{6.11}$$

where, in the case of a white Gaussian noise,

$$\tilde{\mathbf{w}}'_u(\tau_h, \nu'_{D_h}) = k'\mathbf{p}(\tau_h) \circ \mathbf{b}'(\nu'_{D_h}) \tag{6.12}$$

is the matched filter with k' a normalization factor.

Let us now examine the response of the filter if $\tau_h \neq \tau_a$ or $\nu'_{D_h} \neq \nu'_{D_a}$. In this case

$$\begin{aligned} T(\tau_h, \nu'_{D_h}) &= |\tilde{\mathbf{w}}_u'^{\dagger}\tilde{\mathbf{y}}'_{t_u} + \tilde{\mathbf{w}}_u'^{\dagger}\mathbf{n}'|^2 & (6.13)\\ &= |\tilde{\mathbf{w}}_u'^{\dagger}\tilde{\mathbf{y}}'_{t_u}|^2 + |\tilde{\mathbf{w}}_u'^{\dagger}\mathbf{n}'|^2 + 2\mathrm{Re}(\tilde{\mathbf{w}}_u'^{\dagger}\tilde{\mathbf{y}}'_{t_u}\mathbf{n}'^{\dagger}\tilde{\mathbf{w}}'_u). & (6.14) \end{aligned}$$

The first term of this expression corresponds to the noise-free case and can be rewritten as

$$|\tilde{\mathbf{w}}_u'^{\dagger}\tilde{\mathbf{y}}'_{t_u}|^2 = k'^2|\alpha_c|^2|(\mathbf{p}^{\dagger}(\tau_h) \circ \mathbf{b}'^{\dagger}(\nu'_{D_h}))(\mathbf{p}(\tau_a) \circ \mathbf{b}'(\nu'_{D_a}))|^2. \tag{6.15}$$

As $(\mathbf{a} \circ \mathbf{b})^{\dagger}(\mathbf{c} \circ \mathbf{d}) = (\mathbf{a} \circ \mathbf{c}^*)^{\dagger}(\mathbf{b}^* \circ \mathbf{d})$ for vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, and $\mathbf{d}$ with the same length, rewriting the terms not depending on α_c in the previous equation yields

$$k'^2|(\mathbf{p}^{\dagger}(\tau_h) \circ \mathbf{b}'^{\dagger}(\nu'_{D_h}))(\mathbf{p}(\tau_a) \circ \mathbf{b}'(\nu'_{D_a}))|^2 = k'^2|(\mathbf{p}(\tau_h) \circ \mathbf{p}^*(\tau_a))^{\dagger}(\mathbf{b}'^*(\nu'_{D_h}) \circ \mathbf{b}'(\nu'_{D_a}))|^2. \tag{6.16}$$

As the elements of $\mathbf{b}'$ are exponentials of their argument,

$$\mathbf{b}'^*(\nu'_{D_h}) \circ \mathbf{b}'(\nu'_{D_a}) = \mathbf{b}'(\nu'_{D_a} - \nu'_{D_h}) = \mathbf{b}'(\Delta\nu'_D) \tag{6.17}$$

where $\Delta\nu'_D = \nu'_{D_a} - \nu'_{D_h}$. Moreover, let us define $\mathbf{r}'$, the vector of samples of the autocorrelation function of $p(t)$ as

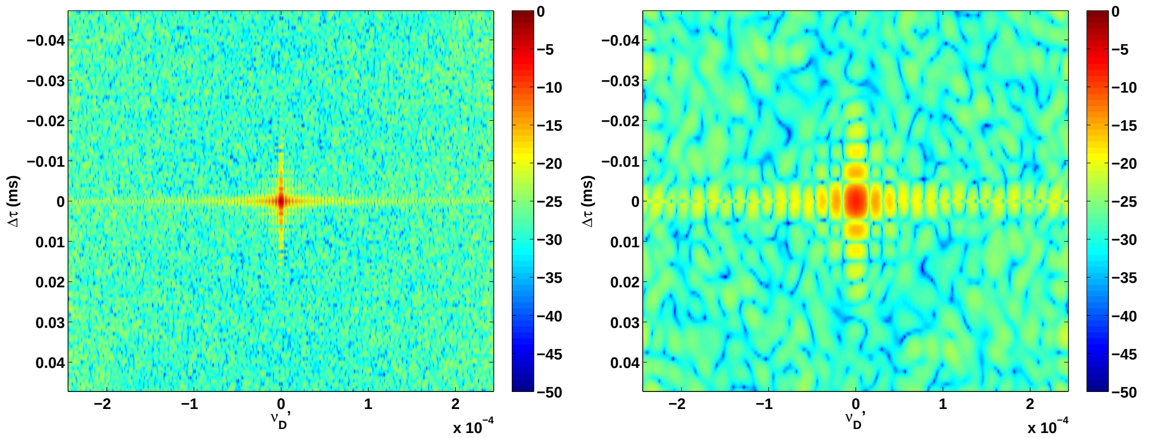
$$\mathbf{r}'(\tau_a - \tau_h, \tau_h) = \mathbf{r}'(\Delta\tau, \tau_h) = \mathbf{p}(\tau_h)^* \circ \mathbf{p}(\tau_a) \tag{6.18}$$

where $\Delta\tau = \tau_a - \tau_h$. Finally, (6.16) becomes

$$\psi(\Delta\tau, \Delta\nu'_D; \tau_h) = k'^2|\mathbf{r}'^{\dagger}(\Delta\tau, \tau_h)\mathbf{b}'(\Delta\nu'_D)|^2. \tag{6.19}$$

A translation of the time origin results in the multiplication of $\mathbf{b}'$ by a constant phase term hence, for $M' \to \infty$, (6.19) becomes independent of τ_h (and of τ_a) and the only dependence on τ_h and τ_a is through their difference, $\Delta\tau$. Let us denote this function by $\psi(\Delta\tau, \Delta\nu'_D)$ and by analogy with its continuous counterpart [146, 170], let us call it the ambiguity function[1].

The ambiguity function is a characteristic of $\mathbf{p}$ and denotes how a signal reflected by a target at time-delay τ_a and with Doppler frequency ν'_{D_a} would be correlated with the signal of a hypothetical target at time-delay τ_h and with Doppler frequency ν'_{D_h}. It gives information on the response of the matched filter $\tilde{\mathbf{w}}'_u$ if it is not matched to the signal to detect. An ideal ambiguity function is maximum at the origin, where the filter is actually matched to the signal to detect and is zero elsewhere. Figure 6.1 (a) and (b) illustrate non ideal ambiguity functions that exhibit



(a) (b)

Figure 6.1: Ambiguity function for a noise-like signal with (a) B = 400kHz, T_{CPI} = 205ms ($M' = 2^{18}$) and (b) B = 100kHz, T_{CPI} = 51ms ($M' = 2^{16}$).

a relatively well pronounced peak at the origin but also strong sidelobes. In Figure 6.1 (b), the sidelobes have been exaggerated for didactic purposes. Although these results are further developed in the next section, intuitively it is clear that the "width" of the peak affects the resolution in time-delay, i.e., in range, and in Doppler frequency, i.e., in velocity. If $p(t)$ exhibits some periodicity, the ambiguity function exhibits the same periodicity along the $\Delta\tau$ axis, which makes the target range determination ambiguous, hence the name. A similar issue arises if $p(t)$ contains harmonics which induce a periodicity of the ambiguity function in the Doppler frequency direction $\Delta\nu'_D$.

6.2.4 Range resolution

Considering a cut at $\Delta\nu'_D = 0$ in the ambiguity function is equivalent to assuming that the Doppler frequency of the target is known. The ambiguity function reduces to the squared ampli-

[1]The naming of ψ is ambiguous [146]. Some authors choose the squared magnitude [115, 170], others choose not to include the squared magnitude operator in the definition [67], others do both [146]. The definition including the squared magnitude is taken here.

tude of the autocorrelation function of $p(t)$ and

$$\psi(\Delta\tau, 0) = k'^2|\mathbf{r}'^\dagger(\Delta\tau)\mathbf{1}|^2 = k'^2|\mathbf{p}^\dagger(\tau_a)\mathbf{p}(\tau_h)|^2. \tag{6.20}$$

The Wiener-Khintchine theorem [80] states that the power spectral density of a wide sense stationary process is the Fourier transform of the corresponding (statistical) autocorrelation function, which provides with the formal link between the range resolution, given by the autocorrelation function of $p(t)$, and the "bandwidth" of $p(t)$.

Thus the range resolution achievable by transmitting a signal with complex envelope $p(t)$ only depends on the autocorrelation function of $p(t)$. This fact constitutes the basis of continuous wave radars [133, 146]. As the spectrum of the waveforms commonly used, e.g., linear frequency modulation waveforms, is nearly rectangular, the autocorrelation function exhibits strong sidelobes, known as range sidelobes. This is illustrated in Figure 6.2 for a signal with

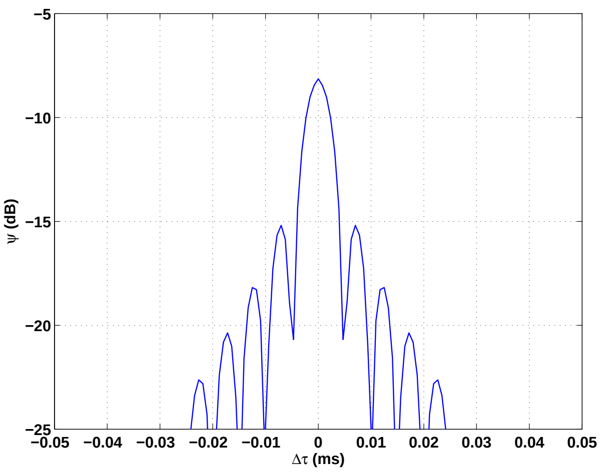


Figure 6.2: Range sidelobes for B = 100kHz.

a bandwidth B of 100kHz. These range sidelobes can be attenuated by shaping the spectrum, i.e., by modifying the transmitted spectrum. Alternatively, a pragmatic approach [115, 146] consists in considering a *mismatched filter*, for which the spectrum shaping is implemented on the receiver, which indeed corresponds to the use of a slightly mismatched filter, hence the name.

6.2.5 Doppler frequency resolution

Considering now a cut in the ambiguity function at $\Delta\tau = 0$, i.e., assuming that the time delay is known, yields

$$\psi(0, \Delta\nu'_D) = k'^2|\mathbf{r}'^\dagger(0)\mathbf{b}'(\Delta\nu'_D)|^2 \tag{6.21}$$

or, if $p(t)$ has a unit magnitude, considering a finite number of samples $M' = T_{\text{CPI}}f_s$ yields

$$\psi(0, \Delta\nu_D) = k'^2|\mathbf{1}^\dagger\mathbf{b}'(\Delta\nu'_D)|^2 \tag{6.22}$$

$$= k'^2\left|\frac{\sin(\pi M'\Delta\nu'_D)}{\sin(\pi\Delta\nu'_D)}\right|^2. \tag{6.23}$$

This function reaches its maximum for $\Delta\nu'_D = 0, \pm 1, \ldots$, and is known as the squared Dirichlet kernel and can be rewritten as

$$\psi(0, \Delta f_D) = k'^2 \left| \frac{\sin(\pi T_{\text{CPI}} \Delta f_D)}{\sin(\pi \Delta f_D / f_s)} \right|^2 . \tag{6.24}$$

with its first zero located at $\Delta f_D = 1/T_{\text{CPI}}$. Hence, in the particular case of constant amplitude complex envelopes, the Doppler frequency resolution only depends on T_{CPI}, i.e., the duration of the observation or the dwell time on the target and *not* on the sampling frequency f_s. The Doppler cut exhibits so-called Doppler sidelobes, due to the finite observation duration. Doppler

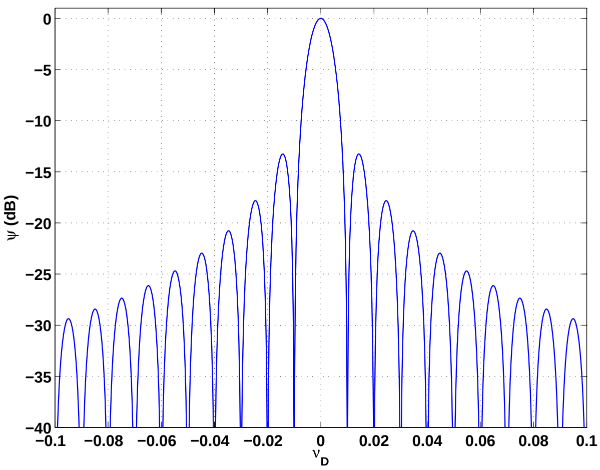


Figure 6.3: Doppler sidelobes for $M' = 100$.

sidelobes are illustrated in Figure 6.3 where, for didactic reasons, a small value of M' is chosen. If a (white-noise) matched filter is used, these sidelobes affect the discrimination of a target in the presence of interfering signals (clutter or other targets).

6.2.6 Fast implementation

The computation of $T(\tau_h, \nu'_{D_h})$ following (6.11) would imply the computation of a filter $\tilde{\mathbf{w}}'_u$ of length M' and the associated scalar product of two vectors of length M', this operation being repeated for each possible τ_h and ν'_{D_h}. As M' is typically very large, these computations are extremely demanding. A possible way to reduce these computational issues is now discussed.

$(\mathbf{a} \circ \mathbf{c})^\dagger \mathbf{d} = \mathbf{a}^\dagger (\mathbf{c}^* \circ \mathbf{d})$ for arbitrary vectors $\mathbf{a}$, $\mathbf{c}$, and $\mathbf{d}$ of equal length. Hence the application of the white-noise matched filter (6.12) to the data

$$\tilde{\mathbf{w}}'^\dagger_u(\tau_h, \nu'_{D_h})\tilde{\mathbf{y}}'_u = k'(\mathbf{p}(\tau_h) \circ \mathbf{b}'(\nu'_{D_h}))^\dagger \tilde{\mathbf{y}}'_u = k'\mathbf{b}'^\dagger(\nu'_{D_h})(\mathbf{p}^*(\tau_h) \circ \tilde{\mathbf{y}}'_u) \tag{6.25}$$

can be interpreted as the scaled Fourier transform of the demodulated data vector $\mathbf{y}'_u = \mathbf{p}^* \circ \tilde{\mathbf{y}}'_u$ since $\mathbf{b}'(\nu'_{D_h})$ is the Fourier eigenfunction.

As noted in [161], the maximum Doppler frequency for actual physical targets is typically much lower than the sampling frequency f_s required to match the modulating signal bandwidth.

Indeed, for example, in the case of a signal with a bandwidth of 200kHz at a carrier frequency of 1GHz and a target velocity of 200km/h, the maximum Doppler frequency $f_{D_{\max}}$ amounts to 370Hz while a sampling frequency f_s between 0.5 and 1MHz is typically considered. The ratio $S = f_s/(2f_{D_{\max}})$ can thus be very high. The demodulated signal $\mathbf{p}^* \circ \tilde{\mathbf{y}}'_u$ can thus be downsampled with a factor S without losing the information about possible targets. In order to avoid aliasing, the noise present at frequencies larger than $f_{D_{\max}}$, a low-pass filter is required prior to subsampling. This filter should have a cut-off frequency equal to $f_{D_{\max}}$. After filtering, the spectral analysis is performed using a filter $\mathbf{b}(\nu_D)$ of length $M = M'/S$ where $\nu_D = f_D/f_{D_{\max}}$ is the modified reduced Doppler frequency. One also has $\nu_D = S\nu'_D$. Note that the vector $\mathbf{b}(\nu_D)$ is exactly the temporal steering vector defined in (3.19). The absence of the apostrophe indicates that the quantities are related to those defined in Chapter 3. The result can thus be written as

$$k\mathbf{b}^\dagger(\nu_{D_h})\mathbf{W}(\mathbf{p}^*(\tau_h) \circ \tilde{\mathbf{y}}'_u) \approx \tilde{\mathbf{w}}'^\dagger_u(\tau_h, \nu'_{D_h})\tilde{\mathbf{y}}'_u \tag{6.26}$$

where $\mathbf{W}$ is the $M \times M'$ matrix implementing the FIR anti-aliasing filter and the subsampling by a factor S. k is a normalization constant.

If a moving average filter of length S is considered, the matrix $\mathbf{W}$ is of the form

$$\mathbf{W} = \frac{1}{S}\begin{pmatrix} \mathbf{1}_S^T & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{1} & \mathbf{0}_S^T & \cdots & \mathbf{0} \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & & \cdots & \mathbf{1}_S^T \end{pmatrix} \tag{6.27}$$

where $\mathbf{1_S}$ is a column vector of length S filled with ones. Although, as noted in [87], a moving average filter does not have ideal characteristics, it allows a very economic implementation [161]. These non-ideal characteristics impact the detectability of targets when their the Doppler frequency approaches $f_{D_{\max}}$.

6.2.7 Discussion

This section only briefly addresses the topic of the ambiguity function and the corresponding estimation of the target range (time-delay) and velocity (Doppler frequency) in the noise-free case.

A more detailed analysis of the achievable accuracy in range and velocity in the presence of a white noise is presented in [170] together with properties, also discussed in [146], of the continuous ambiguity function.

6.3 Single channel and clutter at the range of interest

As an intermediate step, let us now consider the case for which clutter is present but only at the range of interest.

6.3.1 Clutter signal model

The considered clutter model is described in Section 3.5.2 and can be written as

$$\tilde{\mathbf{y}}'_{c_u} = \sum_{k=1}^{K} \alpha_{c,k}\mathbf{p}(\tau_a) \circ \mathbf{b}'(\nu'_{D_a,k}) \tag{6.28}$$

$$= \mathbf{p}(\tau_a) \circ \sum_{k=1}^{K} \alpha_{c,k}\mathbf{b}'(\nu'_{D_a,k}) \tag{6.29}$$

$$= \mathbf{p}(\tau_a) \circ \mathbf{y}'_{c_u} \tag{6.30}$$

where the complex reflectivity coefficient $\alpha_{c,k} = \alpha_k c_k$ includes the geometric terms c_k from the radar equation. Obviously, besides the larger sampling frequency, the clutter signal is the signal, denoted by $\mathbf{y}'_{c_u}$, that would have been obtained, in the hypothetical absence of range ambiguity effects, with a series of rectangular pulses with a PRF equal to f_s, multiplied by the complex envelope $\mathbf{p}(\tau_a)$.

6.3.2 Clutter covariance matrix

Assuming independence of α_k, the clutter covariance matrix, defined as

$$\tilde{\mathbf{R}}'_{c_u} = E\{\tilde{\mathbf{y}}'_{c_u}\tilde{\mathbf{y}}'^{\dagger}_{c_u}\} \tag{6.31}$$

can be rewritten as

$$\tilde{\mathbf{R}}'_{c_u} = \sum_{k=1}^{K} \sigma^2_{\alpha_k} E\{(\mathbf{p}(\tau_a) \circ \mathbf{b}'(\nu'_{D_a,k}))(\mathbf{p}(\tau_a) \circ \mathbf{b}'(\nu'_{D_a,k}))^{\dagger}\} \tag{6.32}$$

where $\sigma^2_{\alpha_k} = E\{|\alpha_{c,k}|^2\}$. As for vectors $\mathbf{a}$ and $\mathbf{c}$ of identical dimensions, $(\mathbf{a} \circ \mathbf{c})(\mathbf{a} \circ \mathbf{c})^{\dagger} = (\mathbf{a}\mathbf{a}^{\dagger}) \circ (\mathbf{c}\mathbf{c}^{\dagger})$, which can be shown by explicitly writing the expression for each element of the vectors, and by assuming a deterministic complex envelope $\mathbf{p}$, the clutter covariance matrix can be rewritten as

$$\tilde{\mathbf{R}}'_{c_u} = \mathbf{P} \circ \mathbf{R}'_{c_u} \tag{6.33}$$

where

$$\mathbf{P} = \mathbf{p}(\tau_a)\mathbf{p}^{\dagger}(\tau_a) \tag{6.34}$$

has the same form as a covariance matrix taper [59, 60], but is a totally different function.

$$\mathbf{R}'_{c_u} = E\{\mathbf{y}'_{c_u}\mathbf{y}'^{\dagger}_{c_u}\} \tag{6.35}$$

is the covariance matrix of $\mathbf{y}'_{c_u}$.

The interference-plus-noise covariance matrix is given by

$$\tilde{\mathbf{R}}'_u = \mathbf{P} \circ \mathbf{R}'_{c_u} + \mathbf{R}'_n \tag{6.36}$$

where $\mathbf{R}'_n$ is the noise covariance matrix which is diagonal as the noise is assumed uncorrelated from sample to sample. Note that, provided that a suitable normalization of $\mathbf{p}$ is used, the diagonal terms of $\mathbf{P}$ are equal to unity and the previous equation can be rewritten as

$$\tilde{\mathbf{R}}'_u = \mathbf{P} \circ (\mathbf{R}'_{c_u} + \mathbf{R}'_n) = \mathbf{P} \circ \mathbf{R}'_u \tag{6.37}$$

where $\mathbf{R}'_u = \mathbf{R}'_{c_u} + \mathbf{R}'_n$ is the clutter-plus-noise covariance matrix that would have been obtained if a pulse train with PRF f_s had been transmitted, neglecting range ambiguities.

6.3.3 Optimum filter

By using the conclusions of Chapter 4, the filter needed to compute the detection statistic is

$$\tilde{\mathbf{w}}'_u = \tilde{k}'\tilde{\mathbf{R}}'^{-1}_u(\mathbf{p}(\tau_h) \circ \mathbf{b}'(\nu'_{D_h})) \tag{6.38}$$

where $\tilde{k}'$ is a normalization factor.

Let us now analyze the term $\tilde{\mathbf{R}}'^{-1}_u$ in (6.38). The covariance matrix $\mathbf{R}'_u$ can be diagonalized and expressed as

$$\mathbf{R}'_u = \sum_{i=1}^{M'} \lambda'_i \mathbf{u}'_i \mathbf{u}'^{\dagger}_i = \mathbf{U}'\mathbf{\Lambda}'\mathbf{U}'^{\dagger} \tag{6.39}$$

where λ'_i is the i-th eigenvalue and $\mathbf{u}'_i$ the associated eigenvector, $\mathbf{U}' = [\mathbf{u}'_1, \mathbf{u}'_2, \ldots, \mathbf{u}'_{M'}]$ and $\mathbf{\Lambda}' = \mathrm{diag}\{\lambda'_1, \lambda'_2, \ldots, \lambda'_{M'}\}$. If the selected eigenvectors $\mathbf{u}'_i$ are orthonormal,

$$\mathbf{R}'^{-1}_u = \sum_{i=1}^{M'} \frac{1}{\lambda'_i} \mathbf{u}'_i \mathbf{u}'^{\dagger}_i = \mathbf{U}'\mathbf{\Lambda}'^{-1}\mathbf{U}'^{\dagger} \tag{6.40}$$

where the inverse is guaranteed to exist since $\mathbf{R}'_n$ is full rank.

By using (6.37), we get

$$\tilde{\mathbf{R}}'_u = \mathbf{P} \circ (\mathbf{U}'\mathbf{\Lambda}'\mathbf{U}'^{\dagger}) \tag{6.41}$$

$$= \mathbf{P} \circ \left(\sum_{i=1}^{M'} \lambda'_i \mathbf{u}'_i \mathbf{u}'^{\dagger}_i \right) \tag{6.42}$$

$$= \sum_{i=1}^{M'} \lambda'_i (\mathbf{p} \circ \mathbf{u}'_i)(\mathbf{p} \circ \mathbf{u}'_i)^{\dagger} \tag{6.43}$$

$$= \tilde{\mathbf{U}}'\mathbf{\Lambda}'\tilde{\mathbf{U}}'^{\dagger} \tag{6.44}$$

where $\tilde{\mathbf{U}}' = [\mathbf{p} \circ \mathbf{u}'_1, \mathbf{p} \circ \mathbf{u}'_2, \ldots, \mathbf{p} \circ \mathbf{u}'_{M'}]$. As the amplitude of the elements of $\mathbf{p}$ is normalized, we have $\mathbf{p}^* \circ \mathbf{p} = \mathbf{1}$ and $\mathbf{p} \circ \mathbf{u}'_i$ still forms an orthonormal basis

$$\tilde{\mathbf{U}}'^{\dagger}\tilde{\mathbf{U}}' = \mathbf{P} \circ \mathbf{I} = \mathbf{I}. \tag{6.45}$$

The inverse of $\tilde{\mathbf{R}}'_u$ is thus given by

$$\tilde{\mathbf{R}}'^{-1}_u = \tilde{\mathbf{U}}'\mathbf{\Lambda}'^{-1}\tilde{\mathbf{U}}'^{\dagger} \tag{6.46}$$

$$= \sum_{i=1}^{M'} \frac{1}{\lambda'_i}(\mathbf{p} \circ \mathbf{u}'_i)(\mathbf{p} \circ \mathbf{u}'_i)^{\dagger} \tag{6.47}$$

$$= \mathbf{P} \circ \sum_{i=1}^{M'} \frac{1}{\lambda'_i}\mathbf{u}'_i\mathbf{u}'^{\dagger}_i \tag{6.48}$$

$$= \mathbf{P} \circ \mathbf{R}'^{-1}_u \tag{6.49}$$

Hence, from (6.47) and (6.38),

$$\tilde{\mathbf{w}}'_u = \tilde{k}' \sum_{i=1}^{M'} \frac{1}{\lambda'_i}(\mathbf{p} \circ \mathbf{u}'_i)(\mathbf{p} \circ \mathbf{u}'_i)^{\dagger}(\mathbf{p} \circ \mathbf{b}') \tag{6.50}$$

or, as $\mathbf{p}^* \circ \mathbf{p} = \mathbf{1}$, $(\mathbf{p} \circ \mathbf{u}'_i)^{\dagger}(\mathbf{p} \circ \mathbf{b}') = \mathbf{u}'^{\dagger}_i\mathbf{b}'$ and

$$\tilde{\mathbf{w}}'_u(\tau_h, \nu'_{D_h}) = \tilde{k}' \sum_{i=1}^{M'} \frac{1}{\lambda'_i}(\mathbf{p} \circ \mathbf{u}'_i)\mathbf{u}'^{\dagger}_i\mathbf{b}' \tag{6.51}$$

$$= \tilde{k}'\mathbf{p} \circ \sum_{i=1}^{M'} \frac{1}{\lambda'_i}\mathbf{u}'_i\mathbf{u}'^{\dagger}_i\mathbf{b}' \tag{6.52}$$

$$= \mathbf{p} \circ (\tilde{k}'\mathbf{R}'^{-1}_u\mathbf{b}') \tag{6.53}$$

$$= \mathbf{p}(\tau_h) \circ \mathbf{w}'_u(\nu'_{D_h}) \tag{6.54}$$

where $\mathbf{w}'_u(\nu'_{D_h}) = \tilde{k}'\mathbf{R}'^{-1}_u\mathbf{b}'(\nu'_{D_h})$ is the optimum filter that would be obtained if no phase modulation is used.

If a filter $\tilde{\mathbf{w}}'_u$ actually matched to the target present in the signal is used, i.e., $\tau_h = \tau_a$ and $\nu'_{D_h} = \nu'_{D_a}$, the operation

$$\tilde{\mathbf{w}}'^{\dagger}_u(\tau_a, \nu'_{D_a})\tilde{\mathbf{y}}'_u = (\mathbf{p}(\tau_a) \circ \mathbf{w}'_u(\nu'_{D_a}))^{\dagger}\tilde{\mathbf{y}}'_u = \mathbf{w}'^{\dagger}_u(\nu'_{D_a})(\mathbf{p}^*(\tau_a) \circ \tilde{\mathbf{y}}'_u) \tag{6.55}$$

can be interpreted as the demodulation of $\tilde{\mathbf{y}}'_u$, i.e., the multiplication of $\tilde{\mathbf{y}}'_u$ with $\mathbf{p}^*$ and the subsequent application of the filter $\mathbf{w}'_u$ that would have been obtained if no modulation ever occurred.

6.3.4 Fast implementation

Section 6.2.6 argued that, as the maximum Doppler frequency of interest is much smaller than the sampling frequency, the demodulated signal $\mathbf{p}^*(\tau_a) \circ \tilde{\mathbf{y}}'_u$ could be low-pass filtered and down-sampled before further processing, which can be written as

$$\mathbf{b}^{\dagger}(\nu_{D_a})\mathbf{W}(\mathbf{p}^*(\tau_a) \circ \tilde{\mathbf{y}}'_u). \tag{6.56}$$

where the matrix $\mathbf{W}$ implements the low-pass filtering and the subsampling operation. As the filtering operation using filter $\mathbf{w}'(\nu'_{D_h})$ can be interpreted as a spectral analysis [80, Chapter 11] with an optimal rejection of the interference-plus-noise, the same reasoning as in Section 6.2.6 can be applied by subsampling the demodulated signal prior to filtering

$$\mathbf{w}_u^\dagger(\nu_{D_h})\mathbf{W}(\mathbf{p}^*(\tau_h) \circ \tilde{\mathbf{y}}'_u) \approx \tilde{\mathbf{w}}_u'^\dagger(\tau_h, \nu'_{D_h})\tilde{\mathbf{y}}'_u \tag{6.57}$$

where $\mathbf{w}_u(\nu_{D_h}) = k\mathbf{R}_u^{-1}\mathbf{b}(\nu_{D_h})$ with k a normalization constant, $\mathbf{b}(\nu_{D_h})$ is the usual temporal steering vector and $\mathbf{R}_u$ is the interference-plus-noise covariance matrix obtained with a pulse-Doppler radar with a PRF equal to $2f_{D_{\max}}$.

6.3.5 Generalized ambiguity function

Let us now generalize the notion of ambiguity function in the case of a filter which is not the white noise matched filter but the fast implementation of the filter matched to the actual interference (6.57).

The response of the filter matched to a target at (τ_h, ν'_{D_h}) if the target is actually at (τ_a, ν'_{D_a}) is

$$k\mathbf{b}^\dagger(\nu_{D_h})\mathbf{R}^{-1}\mathbf{W}(\mathbf{p}^*(\tau_h) \circ \tilde{\mathbf{y}}'_{t_u}(\tau_a, \nu_{D_a})) = k\alpha_c\mathbf{b}^\dagger(\nu_{D_h})\mathbf{R}_u^{-1}\mathbf{W}(\mathbf{p}^*(\tau_h) \circ \mathbf{p}(\tau_a) \circ \mathbf{b}'(\nu'_{D_a})) \tag{6.58}$$

$$= k\alpha_c\mathbf{b}^\dagger(\nu_{D_h})\mathbf{R}_u^{-1}\mathbf{W}(\mathbf{b}'(\nu'_{D_a}) \circ \mathbf{r}'(\Delta\tau, \tau_h)). \tag{6.59}$$

Let us define the generalized[2] ambiguity function $\psi_g(\Delta\tau, \nu_{D_h}, \nu'_{D_a}; \tau_h)$ as

$$\psi_g(\Delta\tau, \nu_{D_h}, \nu'_{D_a}; \tau_h) = |k\mathbf{b}^\dagger(\nu_{D_h})\mathbf{R}_u^{-1}\mathbf{W}(\mathbf{b}'(\nu'_{D_a}) \circ \mathbf{r}'(\Delta\tau, \tau_h))|^2 \tag{6.60}$$

This function is illustrated in Figure 6.4. The considered interference would result from the clutter if neither the transmitter nor the receiver are moving. The effect of the Doppler-adapted filter is to suppress the interference. The properties of this generalized ambiguity function are difficult to assess in all generality due to the presence of the filtering and subsampling matrix $\mathbf{W}$.

Let us now consider a particularly simple matrix $\mathbf{W}$

$$\mathbf{W} = \mathbf{I}_M \otimes [1, 0, \ldots, 0] \tag{6.61}$$

where the row vector $[1, 0, \ldots, 0]$ has a length S. This matrix corresponds to a subsampling in absence of anti-aliasing filtering. With this matrix, one has $\mathbf{W}\mathbf{b}'(\nu'_{D_a}) = \mathbf{b}(\nu'_{D_a}S)$ and a cut at $\Delta\tau = 0$, for which $\mathbf{r}'(0, \tau_h) = \mathbf{1}$, yields

$$\psi_g(0, \nu_{D_h}, \nu'_{D_a}; \tau_h) = |k\mathbf{b}^\dagger(\nu_{D_h})\mathbf{R}_u^{-1}\mathbf{b}(\nu'_{D_a}S)|^2 \tag{6.62}$$

which is an adapted pattern in Doppler frequency which reaches a maximum equal to 1 for $\nu_{D_h} = \nu'_{D_a}S$. This is illustrated in Figure 6.5 (b) where an interference is considered at $\nu_D = 0$

[2]The adjective "generalized" is used in analogy with the usage of the same adjective to qualify the matched filter when colored noise is considered [82].

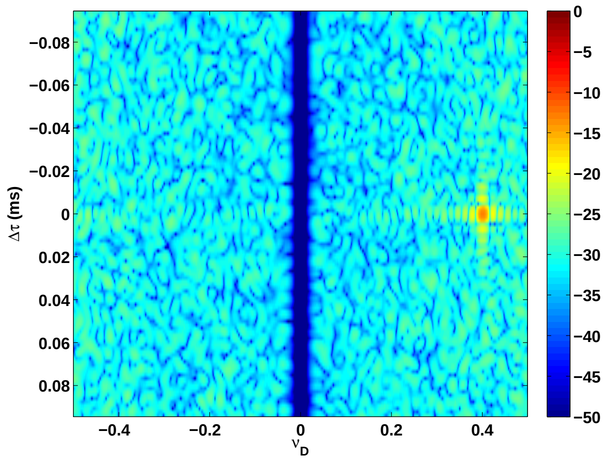


Figure 6.4: Generalized ambiguity function for $\nu_{D_a} = 0.4$, $\tau_a = 0$, $M = 2^{14}$, B = 100kHz and by considering an interference at $\nu_D = 0$ (at all ranges).

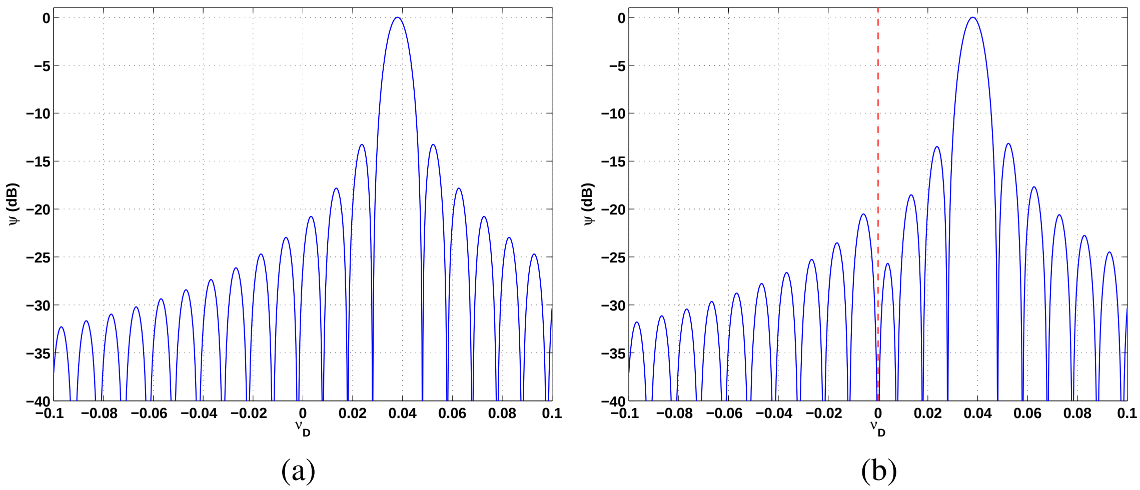


Figure 6.5: Doppler sidelobes for $\nu_{D_a} = 0.038$, for $M = 100$ (a) without interference ($\mathbf{R}_u = \mathbf{I}$) and (b) in the case of an interference at $\nu_d = 0$.

and for $\nu_{D_a} = 0.038$. For didactic reasons, M is taken with an unusual small value. Figure 6.5 (a) presents the pattern that would be obtained by considering $\mathbf{R}_u = \mathbf{I}$, i.e., a white interference. As can be seen, by comparing Figure 6.5 (a) and (b), the pattern in Figure 6.5 (b) is distorted in order to increase the attenuation of the interference.

In order to analyze the behavior of the ambiguity function in the delay direction, let us now assume that the covariance matrix $\mathbf{R}_u$ of the interference is of low-rank. The covariance matrix $\mathbf{R}_u$ can be decomposed as

$$\mathbf{R}_u = \sum_{i=1}^{K} \lambda_i \mathbf{u}_i \mathbf{u}_i^\dagger + \sum_{i=K+1}^{M} \sigma_n^2 \mathbf{u}_i \mathbf{u}_i^\dagger \tag{6.63}$$

where the K first eigenvectors correspond to the interfering signal subspace and the $M-K$ last eigenvalues are equal and correspond to the noise variance σ_n^2, which implies that

$$\mathbf{R}_u^{-1} = \sum_{i=1}^{K} \frac{1}{\lambda_i}\mathbf{u}_i\mathbf{u}_i^\dagger + \sum_{i=K+1}^{M} \frac{1}{\sigma_n^2}\mathbf{u}_i\mathbf{u}_i^\dagger. \tag{6.64}$$

For the purpose of the analysis, let us assume that $\mathbf{b}(\nu_{D_h})$ is "far" from the signal subspace, i.e., $\mathbf{b}^\dagger(\nu_{D_h})\mathbf{u}_i \approx 0$ for $i = 1, \ldots, K$. In this case, the expression $\mathbf{b}^\dagger(\nu_{D_h})\mathbf{R}^{-1}$ reduces to

$$\mathbf{b}^\dagger(\nu_{D_h})\mathbf{R}_u^{-1} \approx \frac{1}{\sigma_n^2}\mathbf{b}^\dagger(\nu_{D_h}). \tag{6.65}$$

Hence for $\nu_{D_h} = \nu'_{D_a} S$,

$$\psi_g(\tau, \nu_{D_h}, \nu_{D_h}/S; \tau_h) \approx \left| \frac{k}{\sigma_n^2}\mathbf{b}^\dagger(\nu_{D_h})\mathbf{W}(\mathbf{b}'(\nu_{D_h}/S) \circ \mathbf{r}'(\Delta\tau, \tau_h)) \right|^2 \tag{6.66}$$

$$= \left| \frac{k}{\sigma_n^2}\mathbf{1}^\dagger\mathbf{W}\mathbf{r}'(\Delta\tau, \tau_h) \right|^2. \tag{6.67}$$

In this case, the generalized ambiguity function, as the usual ambiguity function, reduces to the square magnitude of the autocorrelation function of $p(t)$.

The essential properties of the ambiguity function, derived in Section 6.2, are thus preserved by the generalized ambiguity function.

6.4 Single-channel case with clutter at any range

6.4.1 Clutter signal model

Similarly as in Section 6.3.1, let us consider the clutter model

$$\tilde{\mathbf{y}}'_{c_u} = \sum_{r=1}^{N_r}\sum_{k=1}^{K} \alpha_{c,k,r}\mathbf{p}(\tau_r) \circ \mathbf{b}'(\nu'_{D,k,r}) \tag{6.68}$$

$$= \sum_{r=1}^{N_r} \mathbf{p}(\tau_r) \circ \sum_{k=1}^{K} \alpha_{c,k,r}\mathbf{b}'(\nu'_{D,k,r}) \tag{6.69}$$

$$= \sum_{r=1}^{N_r} \mathbf{p}(\tau_r) \circ \mathbf{y}'_{c_u,r} \tag{6.70}$$

where N_r clutter rings are considered, $\alpha_{c,k,r}$ is the complex reflectivity of the clutter patch k at range ring r, including the geometric terms of the radar equation, τ_r is the time delay corresponding to range ring r, $\nu'_{D,k,r}$ is the reduced Doppler frequency of clutter patch k at range ring r and $\mathbf{y}'_{c_u,r}$ is the hypothetical clutter signal that would have been received from range ring r if a pulse-Doppler radar with PRF f_s would have been used.

6.4.2 Effect of clutter at other range rings

Due to the changing $\mathbf{p}(\tau_r)$ at each range, it is not possible to obtain a closed-form expression for the inverse of the interference (clutter)-plus-noise covariance matrix in this case. Given the huge length of the vector $\tilde{\mathbf{y}}_c$, even computing the interference (clutter)-plus-noise covariance matrix is not feasible.

Let us consider the filter (6.53)

$$\tilde{\mathbf{w}}'_u(\tau_h, \nu'_{D_h}) = k'\mathbf{p}(\tau_h) \circ (\mathbf{R}'^{-1}_u \mathbf{b}'(\nu'_{D_h})) \tag{6.71}$$

which is optimum if only clutter at the range of interest τ_h is present, and let us assess its performance in the current situation.

The way in which the filter responds to clutter at other ranges than τ_h, is given by the generalized ambiguity function ψ_g (6.60).

As ψ_g is typically very small for $\Delta\tau \neq 0$, the contribution of the clutter at the other ranges is greatly attenuated.

6.5 Multi-channel cases

This section extends the previous discussion to multichannel cases.

6.5.1 Signal model

Let us consider a receiver equipped with N receive channels. Let us assume that the bandwidth B of the transmitted waveform and the maximum distance $d_{\max}$ between two receiving elements are small enough such that no bandwidth-induced angle-dependent decorrelation occurs [60]. More specifically, these assumptions require $d_{\max} \ll c/B$. Under this condition, the angle of arrival of a plane wave results in an element-dependent phase shift as discussed in Section 3.3.1.

The measurements made at each element can be ordered in a vector-valued function of length N and written

$$\check{\mathbf{y}}(t) = \mathbf{a}(\vec{\nu}_s)\tilde{y}(t) \tag{6.72}$$

where $\mathbf{a}(\vec{\nu}_s)$ is the spatial steering vector defined in Section 3.4.2, $\vec{\nu}_s$ is the spatial frequency defining the angle of arrival of the plane wave and $y(t)$ is the complex envelope of the received signal (6.4). The signal at each receive element is sampled with sampling frequency f_s as described by (6.5). If (6.10) and (6.72) are combined and the received samples are lexically ordered, the received signal is modeled by

$$\tilde{\mathbf{y}}'(\tau_a, \vec{\nu}_s, \nu'_{D_a}) = \alpha_c \left(\mathbf{p}(\tau_a) \circ \mathbf{b}'(\nu'_{D_a})\right) \otimes \mathbf{a}(\vec{\nu}_s) + \mathbf{n} \tag{6.73}$$

where $\mathbf{n}$ is a white Gaussian noise vector of length NM'. After defining the vector $\mathbf{p}_N$ as

$$\mathbf{p}_N = \mathbf{p} \otimes \mathbf{1}_N, \tag{6.74}$$

equation (6.73) can be rewritten as

$$\tilde{\mathbf{y}}'(\tau_a, \vec{\nu}_s, \nu'_{D_a}) = \alpha_c \, \mathbf{p}_N \circ (\mathbf{b}'(\nu'_{D_a}) \otimes \mathbf{a}(\vec{\nu}_s)) + \mathbf{n}'' \tag{6.75}$$

$$= \alpha_c \, \mathbf{p}_N \circ \mathbf{v}'(\vec{\nu}_s, \nu'_{D_a}) + \mathbf{n}'' \tag{6.76}$$

where the spatio-temporal steering vector $\mathbf{v}'$ is defined as

$$\mathbf{v}'(\vec{\nu}_s, \nu'_{D_a}) = \mathbf{b}'(\nu'_{D_a}) \otimes \mathbf{a}(\vec{\nu}_s) \tag{6.77}$$

by analogy with (3.17). As noted in Section 6.2.2, this vector differs from (3.17) as the temporal sampling frequency is typically much higher than the sampling frequency considered in (3.17). This difference is denoted by the apostrophe.

6.5.2 Optimum filtering

The developments carried out in Section 6.3 naturally extend to the multichannel case. In particular, the interpretation of the optimum filter as a demodulation operation followed by a "usual" spectral analysis is still valid.

The signal model (6.30) readily extends

$$\tilde{\mathbf{y}}'_c = \mathbf{p}_N(\tau_a) \circ \mathbf{y}'_c \tag{6.78}$$

where

$$\mathbf{y}'_c = \sum_{k=1}^{K} \alpha_{c,k} \mathbf{v}'(\vec{\nu}_{s,k}, \nu'_{D_a,k}) \tag{6.79}$$

is the clutter signal that would have been obtained if a pulse-Doppler train with PRF f_s had been transmitted and if range ambiguities are neglected. Therefore, the clutter-plus-noise covariance matrix (6.37) becomes

$$\tilde{\mathbf{R}}' = \mathbf{P}_N \circ \mathbf{R}' \tag{6.80}$$

where $\mathbf{R}'$ can be thought of as the spatio-temporal clutter-plus-noise covariance matrix that would have been obtained if a pulse train with PRF f_s was transmitted, again neglecting range ambiguities. By using a reasoning similar as in Section 6.3.2, it follows that

$$\tilde{\mathbf{R}}'^{-1} = \mathbf{P}_N \circ \mathbf{R}'^{-1} \tag{6.81}$$

if $\mathbf{p}$ has a constant modulus. Hence, the optimum filter is given by

$$\tilde{\mathbf{w}}' = \mathbf{p}_N \circ \mathbf{w}' \tag{6.82}$$

where

$$\mathbf{w}' = k'\mathbf{R}'^{-1}\mathbf{v}' \tag{6.83}$$

can be interpreted as the spatio-temporal optimum filter that would be obtained if the transmitted signal is a coherent pulse train with PRF f_s and k' a normalization factor. Hence,

$$\tilde{\mathbf{w}}'^{\dagger}\tilde{\mathbf{y}}' = (\mathbf{p}_N \circ \mathbf{w}')^{\dagger}\tilde{\mathbf{y}}' = \mathbf{w}'^{\dagger}(\mathbf{p}_N^* \circ \tilde{\mathbf{y}}') \tag{6.84}$$

and optimum filtering can still be seen as a demodulation step followed by the usual optimum space-time filtering.

6.5.3 Fast implementation

As in Section 6.3.4, which found that the demodulated signal $\mathbf{p}^* \circ \mathbf{y}'_1$ could be low-pass filtered and subsampled, as the maximum Doppler frequency of interest is much smaller than the sampling frequency f_s, the signal from each receive channel can be low-pass filtered and subsampled, before performing the spatio-temporal filtering

$$\tilde{\mathbf{w}}'^{\dagger}\tilde{\mathbf{y}}' \approx \mathbf{w}^{\dagger}(\vec{\nu}_s, \nu_{D_h})\mathbf{W}_N(\mathbf{p}_N^* \circ \tilde{\mathbf{y}}') \tag{6.85}$$

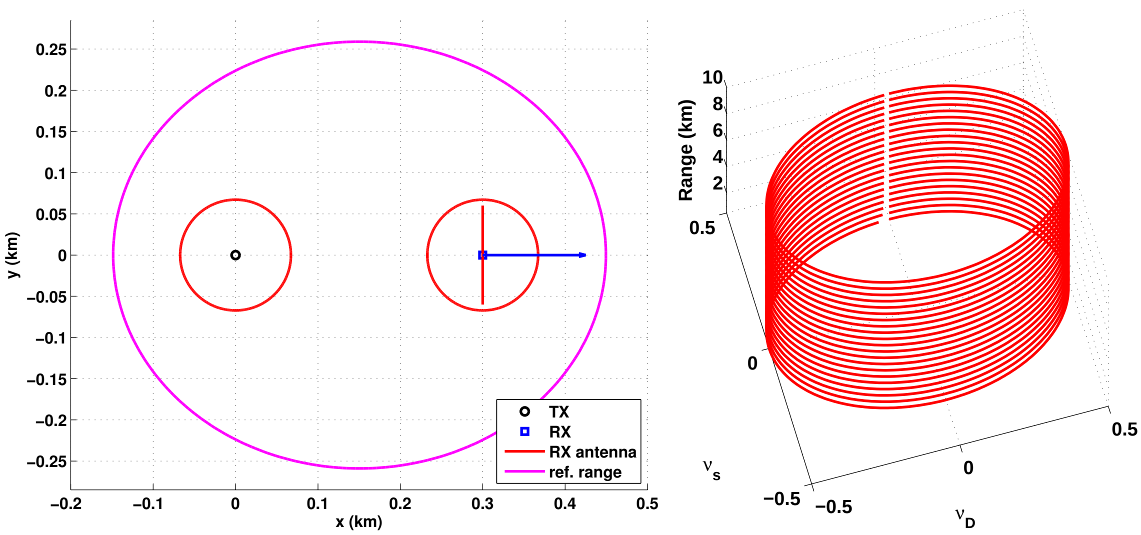
where $\mathbf{W}_N$ is a $NM \times NM'$ matrix implementing the temporal averaging and sampling, independently for each channel. The subscript N denotes an extension of the temporal averaging and sampling matrix $\mathbf{W}$ used in the single channel case

$$\mathbf{W}_N = \mathbf{W} \otimes \mathbf{I}_N. \tag{6.86}$$

$\mathbf{w}$ is the spatio-temporal optimum filter which would be obtained if a pulsed waveform had been considered with PRF f_s/S where S is the subsampling factor.

6.5.4 Illustration

In order to illustrate the developments of the previous section, let us now present an example on simulated data. The considered scenario is depicted in Figure 6.6 (a) and involves a static GSM base station and a receiver located on the ground, moving at a speed of 10m/s. The receiving array is a $\lambda/2$-spaced 8-element forward-looking uniform linear array (ULA). As a ULA is



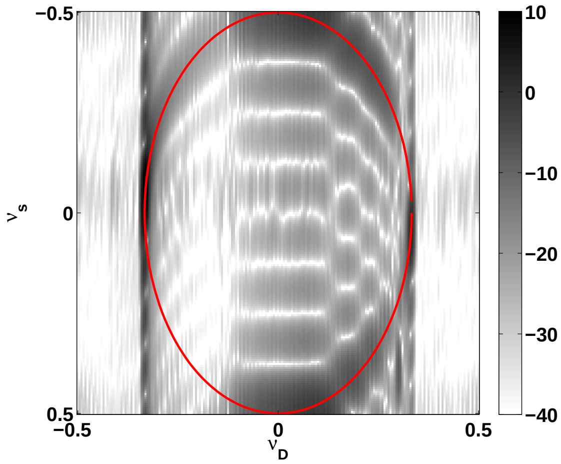
(a) (b)

Figure 6.6: (a) Considered scenario with the isorange considered drawn as a solid line (b) Clutter power spectrum locus in function of the range.

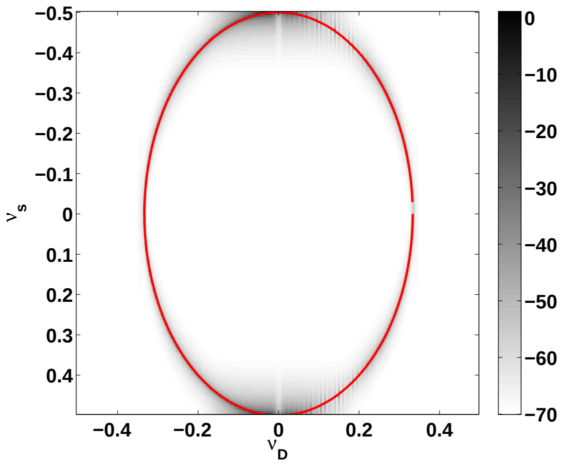
considered, only one spatial frequency can be measured. This spatial frequency is denoted ν_s.

The transmit and the receive antennas have an omnidirectional radiation pattern to exacerbate the influence of the clutter. However, in practice, the radiation pattern of GSM base-stations is far from being omnidirectional. Bistatic scenarios generally involve a geometry-induced range-dependence of the clutter statistic. However, in the particular case of a static transmitter and when the receiver is located on the ground, the clutter statistic does not exhibit any geometry-induced range-dependence as discussed in Chapter 5. Figure 6.6 (b) depicts the clutter power spectrum locus for different ranges.

Figure 6.7 (a) depicts the power spectrum of the subsampled demodulated signal $\mathbf{W}_N(\mathbf{p}_N \circ \tilde{\mathbf{y}}')$ at the range of interest. The clutter power spectrum locus is plotted as a thin line. The contribution of the clutter along the clutter power spectrum locus is clearly visible.



(a)



(b)

Figure 6.7: (a) Power spectrum of the demodulated signal and clutter power spectrum locus, (b) Power spectrum of the clutter covariance matrix.

A modeled interference-plus-noise covariance matrix is considered for the computation of the optimum filter $\mathbf{w}$. The power spectrum of the considered covariance matrix is illustrated in Figure 6.7 (b). Again, the power spectrum of the covariance matrix is located along the clutter power spectrum locus.

The test statistic $T(\nu_s, \nu_D)$ resulting from the application of the optimum filter to the demodulated signal $\mathbf{p}_N \circ \tilde{\mathbf{y}}'$ is illustrated in Figure 6.8. The thin solid line corresponds to the clutter power spectrum locus. As can be seen, the clutter contribution is filtered out, leaving the target standing out at $(\nu_s, \nu_D) = (0.4, 0.3)$.

6.6 Feasibility of STAP with signals of opportunity

6.6.1 Introduction

The previous sections have developed the theory behind space-time processing by using signals of opportunity. Let us now present results from actual measurements.

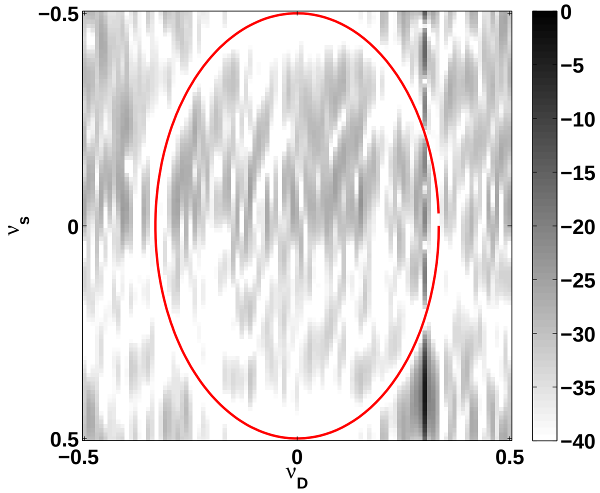


Figure 6.8: Adapted matched filter output and clutter power spectrum locus.

6.6.2 Signal acquisition and pre-processing

A block diagram of a passive GSM-based radar (receiver) is depicted in Figure 6.9. A two-channel receiver is used, the two antennas being arranged to form an array. The array is oriented such that the broadside direction is pointing to the targets. After amplification by the low noise

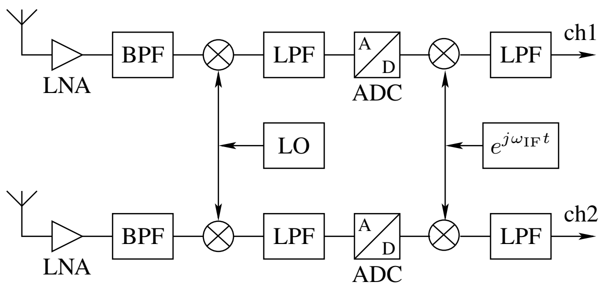


Figure 6.9: Block diagram of the receiver.

amplifier (LNA) and filtering by the band pass filter (BPF) to keep only the GSM downlink band, the signal is down-converted to intermediate frequency and sampled. Once sampled, the received signal is further down-converted by using digital down-conversion (DDC). Performing the latter down-conversion step digitally eliminates imbalances between the in-phase and quadrature channels.

To correct possible asymmetries between the two channels, a calibration step is required. The calibration is also used to measure the phase center of the antennas in order to extract correct direction information from the measurements.

As the bandwidth of the receiver is much larger than that of one GSM channel, the signals from several GSM base stations can be received at once. A typical spectrum of the baseband acquired signal is given in Figure 6.10, where different GSM downlink channels are clearly

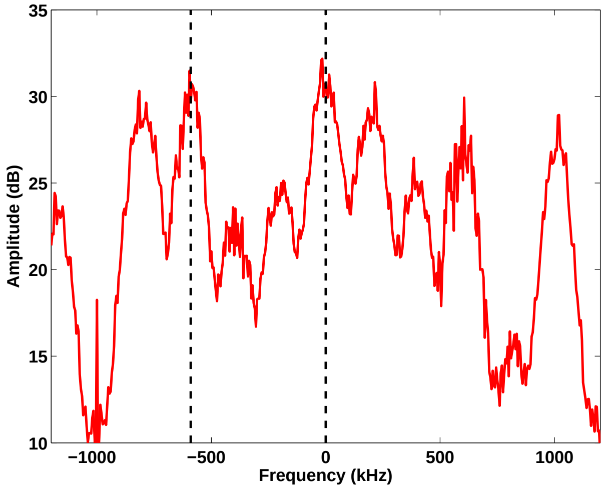


Figure 6.10: Spectrum of the received signal

visible. The two most powerful channels, located around 0Hz and -600kHz, correspond to two different base stations.

The received signal contains the direct signal coming from the GSM base-station transmitter and the echoes of the GSM base-station signal backscattered by the vegetation, buildings, and targets. To be able to perform the coherent processing, it is fundamental to know the reference signal $\mathbf{p}$ broadcast by the GSM base-station. Reference [87] describes a method to blindly extract the reference signal from an array of sensors using adaptive beamforming. Indeed, to avoid artifacts, it is essential that the reference signal does not contain any echo from any target. If the reference signal contained echoes from targets, the concerned targets would be attenuated by the echo cancellation processing [87].

6.6.3 End-to-end results on measured data

The results presented here correspond to real measurements described in more details in [87, 126]. The geometric configuration of the transmitter, receiver, and target is illustrated in Figure 6.11. The receiving antenna array is static and has 2 elements separated by about 0.8λ. The scenario involves a cooperative vehicle (a small van) approaching the receiver and yielding a Doppler frequency of about -50Hz. With this Doppler frequency and the achieved frequency resolution, the vehicle signature is hidden in the sidelobes of the (untapered) matched filter $\tilde{\mathbf{w}} = \mathbf{p}_2 \circ \mathbf{v}$. A tapered matched filter (TMF) which a Hamming-window tapering yields the classical angle-Doppler diagram of Figure 6.12 (a) where the clutter signal is responsible for the large response around $\nu_D = 0$. By using the proposed STAP approach, the contributions of the clutter (including both the static part and the small internal clutter motion (ICM) components) can be removed, leaving the target echo standing out as can be seen at $(\nu_s, \nu_D) = (0.35, -0.07)$

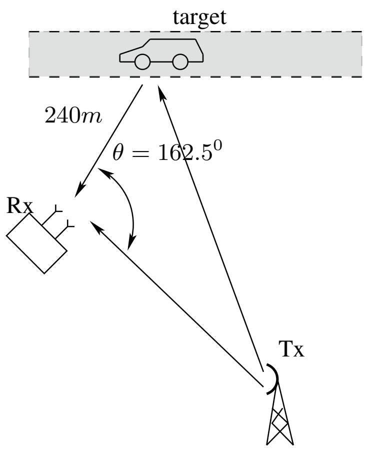


Figure 6.11: Geometric configuration for the real measurement.

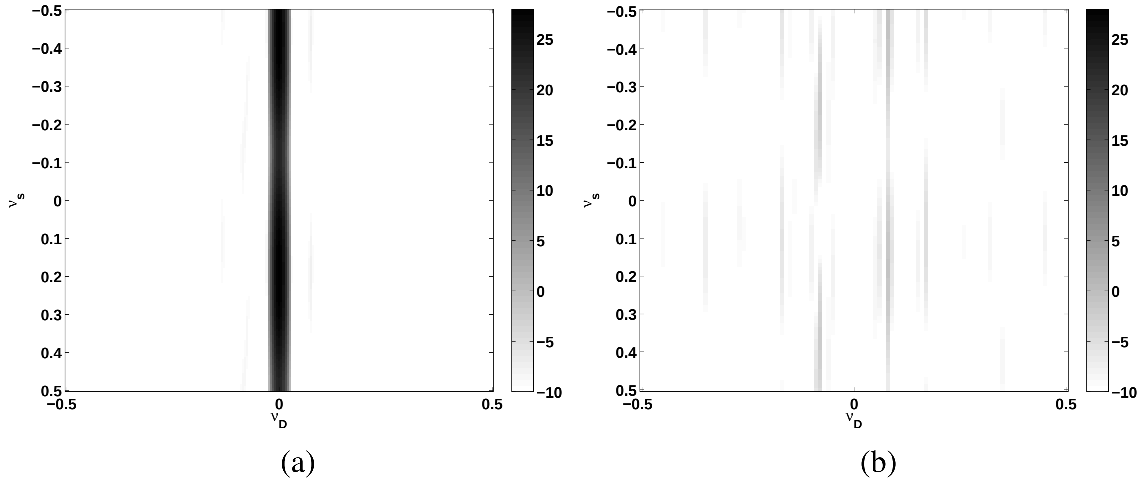
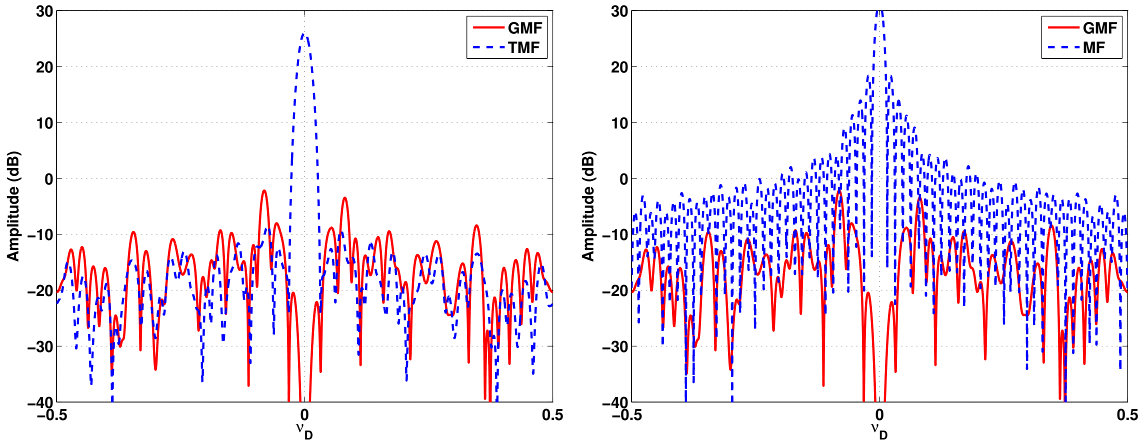


Figure 6.12: (a) White-noise matched filter. (b) Generalized matched filter.

(these values correspond to one of the ambiguous solutions) in Figure 6.12 (b); the other signatures are due either to reflexions or to other (uncooperative) targets. The generalized matched filter (GMF) is obtained from (6.85) using a modeled covariance matrix involving ICM.

Figure 6.13 (a) presents a cut along $\nu_s = 0.35$ in the angle-Doppler diagrams presented in Figure 6.12. Clearly, the generalized matched filter essentially removes the components due to the clutter, located around zero-Doppler. The amplitude of the signature of the vehicle at $\nu_D = -0.07$ is smaller for the tapered matched filter than for the generalized matched filter, which is due to the tapering losses. Figure 6.13 (b) shows a similar cut, in the case of a non-tapered white-noise matched filter (MF). Clearly, the Doppler sidelobes cause “leakage” from the clutter in the target Doppler cell, making detection impossible.

Note that the clutter angle-Doppler diagram does not exhibit the classical space-time coupling as the transmitter and the receiver are both fixed. Hence space-time processing, in this particular scenario, is not really required and a temporal processing would be sufficient. Nevertheless, this



(a) (b)

Figure 6.13: (a) Hamming window. (b) Rectangular window.

case can be seen as a degenerate case of joint space-time processing.

6.7 Other approaches

Generalized matched filtering approaches do not take into account ambiguities other than at the origin $\Delta\tau = 0$. Ambiguities at other ranges, i.e., for $\Delta\tau \neq 0$, may cause the clutter signal to "leak" and hide possible targets.

If the clutter signature is at zero Doppler, i.e., if the transmitter and the receiver are fixed, the classical approach consists in performing some form of echo cancellation [87, 88, 89]. These approaches subtract from the received signal a sum of delayed and possibly phase-shifted versions of the reference signal $\mathbf{p}$. These approaches fail if the clutter exhibits an angle-Doppler dependence since no provision is made to account for a Doppler shift.

The CLEAN algorithm, initially used in radio astronomy, exploits the fact that the targets of interest can be considered as point targets. This algorithm basically consists in iteratively subtracting the scaled (and possibly phase-shifted) point target response of the system at the location of the brightest residual. This method can be applied to radar target detection [30, 44], where the point target response is the ambiguity function of the transmitted waveform, and permits the detection of closely-located targets with large reflectivity differences. The CLEAN algorithm can also be applied to remove clutter [98, 99, 100] if the clutter exhibits an angle-Doppler dependence.

Reference [87] propose an extension of the CLEAN algorithm to handle target detection using signals of opportunity. The algorithm can be interpreted as the decomposition of the received signal in a basis of non-orthogonal functions. Each of these basis functions consists in a time-delayed, angle-Doppler shifted version of the reference signal. The decomposition, which is not unique, is performed by a pursuit-like algorithm [104], where only a fraction of the amplitude of the mode is subtracted at each iteration. This can be seen as an extension of the echo-cancellation methods applied to non-zero Doppler shifts.

Clearly, as the contributions due to each point source are successively subtracted from the received signal, the "leakage" due to ambiguities of the waveform disappears. Eventually, all point-targets (extended targets can be modeled as a group of point targets) are removed from the signal, in which case only noise remains. The difficulty lies in sorting the point targets into actual targets and clutter, which, for instance, can be done by assuming that point targets "close" to the theoretical clutter power spectrum locus belong to the clutter while the other point targets are actual targets.

6.8 Extension to multistatic configurations

The use of transmitters of opportunity naturally extends to multistatic configurations. Obviously, several operating transmitters are needed and their signals must be separable from each other. As shown in Figure 6.10, this is clearly the case for GSM base stations: the signals from the different transmitters are separated in frequency and each transmitter forms with the receiver a bistatic radar. Note that the different transmitters of some digital television (DVB-T) networks that take advantage of the resilience of the modulation used to operate all the transmitters at the same frequency, can be difficult to separate from each other. This is for instance the case of one of the DVB-T networks in Belgium.

The ambiguity function can be defined for each bistatic radar. In the case of a combination of bistatic radars, it may prove useful to rescale the axes of the ambiguity function in order to have the ambiguity function displayed in actual quantities, i.e., velocity and position, as done in [20, 135, 167]. This approach gives a clear view of the positioning accuracy that can be achieved.

Chapter 7

Clutter map estimation

7.1 Introduction

This chapter first reviews the issue of the estimation of the interference-plus-noise covariance matrix together with the solutions presented in the literature. It next presents a maximum likelihood estimation method based on the available information about the covariance matrix structure and formulates the method as the estimation of the clutter map also called scattering function. It develops the method in the case of the estimation of the interference-plus-noise covariance matrix of a bistatic radar and discusses the extension to the estimation of the interference-plus-noise covariance matrix of a multistatic radar, i.e., whether information can be gained from the overlap of the clutter maps of the different bistatic radars.

7.2 Interference-plus-noise covariance matrix estimation

Let us first consider a bistatic radar part of a multistatic radar. In order to make the notation clear, the subscript p denoting the bistatic radar of interest is dropped. Section 7.9 discusses the feasibility of the extension to multistatic radars.

As discussed in Chapter 4, the optimal detection filter is given by

$$\hat{\mathbf{w}} = k\hat{\mathbf{R}}^{-1}\mathbf{s} \tag{7.1}$$

where $\hat{\mathbf{R}}$ is the maximum likelihood estimate (MLE) of the interference-plus-noise covariance matrix.

The estimation of the interference-plus-noise covariance matrix is a central issue in STAP [60, 85], which indeed makes the processing algorithm adaptive. Given K samples of jointly Gaussian, independent, and identically distributed target-free interference-plus-noise data vectors $\mathbf{y}_{i+n_k}$, $k = 1, \ldots, K$ and if no other *a priori* information is available, it is well known [22, 54, 60, 171] that the MLE of the interference-plus-noise covariance matrix is the sample covariance matrix (SCM) of the available data vectors

$$\hat{\mathbf{R}} = \frac{1}{K}\sum_{k=1}^{K}\mathbf{y}_{i+n_k}\mathbf{y}_{i+n_k}^{\dagger}. \tag{7.2}$$

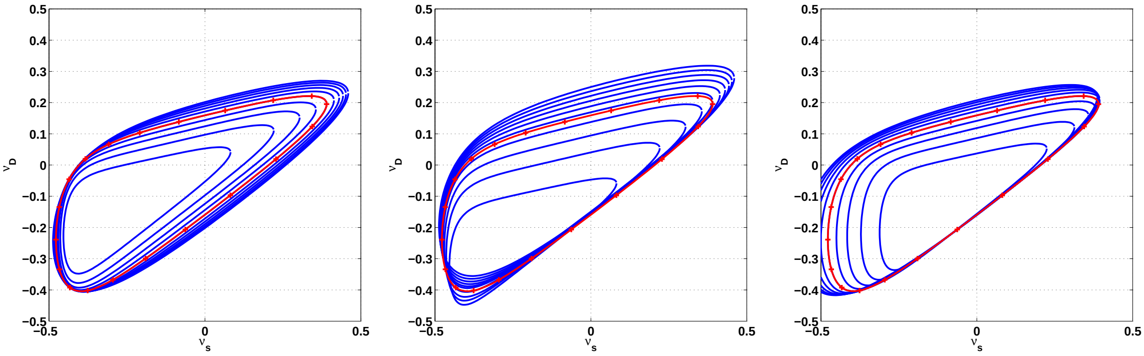
K should be larger than $2NM$, with N the number of antenna elements and M the number of pulses, to obtain a mean performance within 3dB of the optimum performance [142]. One of the key issues in STAP is that a large number of independent and identically distributed interference-plus-noise data vectors are not available.

Independent interference-plus-noise data vectors are usually obtained by considering data vectors at ranges gates around the range gate of interest. The direct application of this approach leads to the use of the sample covariance matrix as an estimate of the covariance matrix. The use of the sample covariance matrix in STAP is known as the sample matrix inversion (SMI) [142] and the modified sample matrix inversion (MSMI) [114].

Typically, a very large number of independent samples are required to achieve acceptable performance. The geographical range over which the samples are taken can thus be extremely large. As the interference-plus-noise data vectors consist in clutter contribution plus noise, they largely depend on the ground cover at the ranges corresponding to the recorded data vectors. Obviously, it is not reasonable to expect that the ground cover be homogeneous over large areas with a few anecdotic exceptions. Mixed land-sea interfaces are an example of situations where large differences in ground scattering coefficients are expected. Similarly, targets or clutter discretes such as a large building also violate the homogeneity assumption. In order to cope with inhomogeneities, a possible solution consists in excluding the inhomogeneous data vectors from the set used to estimate the covariance matrix [113, 114].

Moreover, the structure of the covariance matrix strongly depends on the particular scenario of interest. For monostatic non side-looking and most bistatic observation geometries, the probability density function (PDF) of the clutter data vector is range dependent due to the range dependence of the clutter power spectrum locus as shown in Chapter 5. The evolution of the clutter power spectrum locus as a function of the range is illustrated in Figure 7.1 (a). A possible solution to cope with this range-dependence consists in applying some sort of range-dependence compensation, by which the interference-plus-noise data vectors are transformed to superpose the clutter power spectrum locus at the range of the data with the clutter power spectrum locus at the range of interest. [19, 137] propose a method which achieves a partial superposition by shifting the measurements in Doppler frequency. The superposition of the clutter power spectrum loci is illustrated in Figure 7.1 (b) where perfect superposition is only achieved at one particular point of the spectrum. [71, 73, 111] extend this method in order to align the maxima of the clutter power spectrum. The resulting superposition is illustrated in Figure 7.1 (c), where the clutter power spectrum locus at the different ranges are only partially superposed. [92, 93, 94, 95] propose a method to perform a full power spectrum locus alignment at all ranges with a perfect alignment.

The range-dependence of the clutter statistics can also be taken into account by considering a range-varying filter [86, 109, 110, 176, 177], which only keeps the two first terms of its expansion in a Taylor series, hence doubling the size of the problem. However, these methods, with the



(a) (b) (c)

Figure 7.1: Power spectrum locus (a) no superposition, (b) superposition achievable using Doppler frequency shifts only, (c) superposition achievable using Doppler frequency shift and spatial frequency shift.

exception of the full power spectrum locus alignment method, typically fail if the clutter exhibits pronounced range-dependence characteristics, which is common in bistatic radars.

Another possible solution to cope with the small number of available independent and identically distributed interference-plus-noise data vectors consists in reducing the dimension of the problem [61, 84, 85] by applying data-independent transformations to project space-time data to a lower-dimensional subspace. The transformed interference-plus-noise covariance matrix, of a smaller size, can then be estimated using a smaller number of transformed data vectors. Similarly, the reduced-rank methods [62, 64] consist in keeping only the eigenvectors corresponding to the dominant eigenvalues of the estimated interference-plus-noise covariance matrix. Indeed, these large eigenvalues are relatively accurately estimated by using a low number of interference-plus-noise data vectors. The other eigenvalues are synthesized. Signal-dependent rank reduction methods are also proposed in [50, 51, 72, 129]. A taxonomy of the different STAP methods can be found in [55].

Finally, there is a family of methods that exploits the fact that the covariance matrix actually has a particular structure. Indeed, it is well known that the clutter covariance matrix exhibits a Toeplitz-block-Toeplitz structure if the receiver is equipped with a uniform linear array (ULA) and if a constant PRF is used. This property is exploited in [13, 46, 102, 103, 130, 168] to ease the estimation of the covariance matrix. The covariance matrix still exhibits a particular structure, even in the non-ULA case, which can be used [23, 25] to further constrain the estimation and hence reduce the sample support. More subtle structures such as the modeling of decorrelation effects as covariance matrix tapers can also be exploited [59, 61].

7.3 Problem statement

7.3.1 Modeling and notations

Chapter 3 gives the model of the interference-plus-noise signal

$$\mathbf{y} = \sum_{i=1}^{K} \alpha_i \mathbf{s}_i + \mathbf{n} \tag{7.3}$$

where the $i+n$ subscript is dropped in order to make the notations clearer, K is the number of clutter patches approximating the integral, α_i is the complex reflectivity of clutter patch i and is assumed to be a zero mean circular complex Gaussian random variable with unknown variance r_{α_i}. Furthermore, the complex reflectivity α_i is assumed uncorrelated from clutter patch to clutter patch. $\mathbf{s}_i$ is the steering vector, of length NM, corresponding to the clutter patch i and includes a scale factor that takes into account the geometric factors of the radar equation (range attenuation, element radiation pattern, ...). $\mathbf{n}$ denotes the thermal noise, assumed circular complex Gaussian with covariance matrix $\mathbf{R}_n$. Usually, $\mathbf{R}_n = \sigma_n^2 \mathbf{I}$. $\mathbf{R}_n$ is assumed known as it can easily be obtained by direct measurement, for instance during the interval between pulse emission and the reception of the corresponding ground echoes as described in [96]. Equation (7.3) can be rewritten in matrix form as

$$\mathbf{y} = \mathbf{S}\boldsymbol{\alpha} + \mathbf{n} \tag{7.4}$$

where $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_K]^T$ and $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_K]$.

By using this model, the interference-plus-noise covariance matrix $\mathbf{R} = E\{\mathbf{y}\mathbf{y}^\dagger\}$ is

$$\mathbf{R} = \mathbf{S}\mathbf{R}_\alpha \mathbf{S}^\dagger + \mathbf{R}_\mathbf{n} \tag{7.5}$$

where $\mathbf{R}_\alpha = \mathrm{diag}\{\mathbf{r}_\alpha\}$ with $\mathbf{r}_\alpha = \{r_{\alpha_1}, r_{\alpha_2}, \ldots, r_{\alpha_K}\}^T$ and $r_{\alpha_i} = E\{\alpha_i \alpha_i^*\}$.

7.3.2 On the use of the *a priori* information

The range-dependence compensation methods reviewed in Section 7.2 consist in applying a transform to the data vector in order to compensate for the difference in clutter power spectrum locus between the range of the data vector and the range of interest. Note that some methods [19, 71, 73] provide with an explicit transformation while other methods [92] provide with a procedural method. Once the range dependence in the data vectors is compensated for, these vectors are assumed to have the same statistical distribution (i.e., the same covariance matrix) as the clutter at the range of interest and equation (7.2) is then used to obtain an estimate of the covariance matrix. These compensation methods are implicitly justified by the assumption that the maximum likelihood property of the sample covariance matrix (7.2) still holds even if the transformed vectors are used instead of $\mathbf{y}_{i+n_k}$.

Actually, *a priori* knowledge about the form of the evolution of the clutter power spectrum locus as a function of the range is explicitly used in these methods. However, there is no evidence that this *a priori* knowledge is optimally used. Let us also mention that the sample covariance matrix is the MLE of the covariance matrix if no *a priori* information is available [22, 54, 60, 171].

7.3.3 Spatial homogeneity

Another assumption with respect to the spatial homogeneity of the ground clutter is implicitly made when applying most methods described in Section 7.2. These methods consist in transforming the data vector $\mathbf{y}_r$ at range ring r to $\mathbf{y}'_r$ such that the clutter power spectra overlap and in an averaging the external product of the transformed data vector. Indeed, at each range ring r

$$\mathbf{y}_r = \sum_{i=1}^{K} \alpha_{i,r}\mathbf{s}_{i,r} + \mathbf{n} \tag{7.6}$$

where $\alpha_{i,r}$ is the complex amplitude of the signal scattered by clutter patch i at range ring r and $\mathbf{s}_{i,r}$ is the steering vector corresponding to clutter patch i at range ring r. After the ideal transformation,

$$\mathbf{y}'_r = \sum_{i=1}^{K} \alpha_{i,r}\mathbf{s}_i + \mathbf{n} \tag{7.7}$$

where $\mathbf{s}_i$ is the steering vector at the range of interest for clutter patch i. Note that the transformation of $\mathbf{y}_r$ into $\mathbf{y}'_r$, assumed perfect, has for effect that the $\mathbf{s}_i$'s in the expression above are identical for all r. The estimated interference-plus-noise covariance matrix is thus

$$\hat{\mathbf{R}} = \frac{1}{N_r}\sum_{r=1}^{N_r} \mathbf{y}'_r\mathbf{y}'^{\dagger}_r \tag{7.8}$$

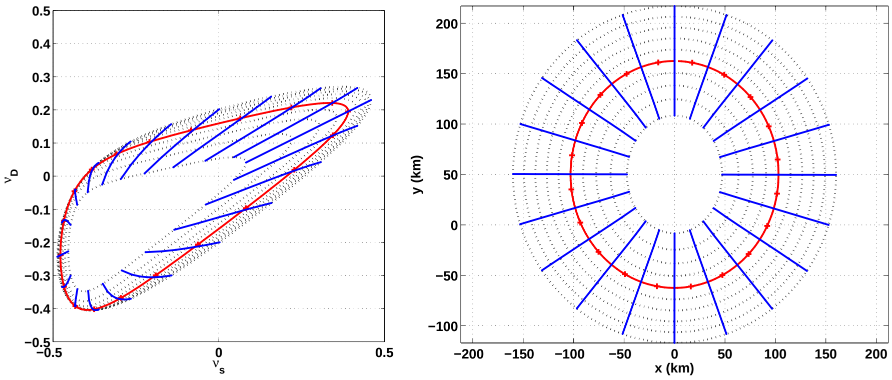
where N_r is the number of range rings considered. Hence

$$\begin{aligned}
E\{\hat{\mathbf{R}}\} &= \frac{1}{N_r}\sum_{r=1}^{N_r}\sum_{i=1}^{K} E\{\alpha_{i,r}\alpha^*_{i,r}\}\mathbf{s}_i\mathbf{s}_i^{\dagger} + \mathbf{R}_n && (7.9)\\
&= \frac{1}{N_r}\sum_{r=1}^{N_r}\sum_{i=1}^{K} r_{\alpha_{i,r}}\mathbf{s}_i\mathbf{s}_i^{\dagger} + \mathbf{R}_n && (7.10)\\
&= \sum_{i=1}^{K} \check{r}_{\alpha_i}\mathbf{s}_i\mathbf{s}_i^{\dagger} + \mathbf{R}_n && (7.11)
\end{aligned}$$

where $r_{\alpha_{i,r}} = E\{\alpha_{i,r}\alpha^*_{i,r}\}$ is the scattering coefficient of clutter patch i at range r and

$$\check{r}_{\alpha_i} = \frac{1}{N_r}\sum_{r=1}^{N_r} r_{\alpha_{i,r}} \tag{7.12}$$

is the estimated scattering coefficient of clutter patch i at the range of interest. This result can be seen as averaging the scattering coefficients along flow lines linking “matching” points along the clutter power spectrum loci as illustrated in Figure 7.2 (a) for a particular bistatic scenario. As there is a direct correspondence between the spectral domain and the spatial domain, this result is equivalent to averaging the scattering coefficients along spatial flow lines on the ground,



(a) (b)

Figure 7.2: The radial lines are the “Flow lines” along which the scattering coefficients are averaged (a) in the spectral domain and (b) in the spatial domain for a particular bistatic scenario. The closed curves are respectively the isorange and the image of the isorange. The red curve in denotes the isorange (and its image) at the range of interest.

as illustrated in Figure 7.2 (b). It implicitly corresponds to assuming that the scattering coefficient is constant along these flow lines. Although this approach is less restrictive than assuming a complete homogeneity of the ground cover, it is nevertheless quite unrealistic. Obviously, this assumption is also made in the cases for which a straightforward application of the sample covariance matrix is used even if the clutter power spectrum locus is not range-dependent.

7.3.4 Structured covariance matrix estimation

Estimation of the clutter covariance matrix assuming a known clutter power spectrum locus can be recast as a structured covariance matrix estimation problem, also called parametric spectral estimation [171]. Hereunder, this chapter presents a maximum likelihood clutter covariance matrix estimation method which explicitly takes into account the known covariance matrix structure. This method implies the resolution of a non-linear equation, generally solved by using the expectation-maximization algorithm [37, 134]. A less computationally demanding alternative is proposed, and, as a comparison, other related methods discussed in the literature are munitioned.

Expressing the estimation problem as a structured covariance matrix estimation problem has two advantages

- The true maximum likelihood estimate of the clutter covariance matrix can be estimated, by taking into account the known structure, which guarantees an optimal use of that knowledge.
- The spatial averaging, intuitively necessary, is naturally introduced as a regularization of the solution. Moreover, as a physical meaning can be given to the computed values, i.e., the scattering coefficients, a physically meaningful spatial averaging can be made, both in range (as is classically done) and in azimuth.

7.3.5 Possible approaches

By assuming a perfect knowledge of $\mathbf{S}$, the unknown parameters in the parametric expression of the interference-plus-noise covariance matrix (7.5) are the scattering coefficients $\mathbf{r}_\alpha$ which need to be estimated.

A first approach [75, 112] consists in estimating $\boldsymbol{\alpha}$ from the data and in using several of these estimates to estimate $\mathbf{r}_\alpha$. As several independent realizations are rarely available, a spatial averaging can be performed, which is detailed further in Section 7.4.

Another approach, more common in other fields like image reconstruction in astronomy [91] or in radar imaging [90, 117, 156, 157], consists in directly estimating the clutter scattering coefficients $\mathbf{r}_\alpha$ from the available data $\mathbf{y}$. The MLE of $\mathbf{r}_\alpha$ is by definition obtained by maximizing the likelihood of $\mathbf{R}_\alpha$. This approach is also called spectral estimation and fully exploits the structure of the covariance matrix of $\mathbf{y}$. It is detailed in Section 7.5.

Once these coefficients are obtained, an estimate of the covariance matrix $\mathbf{R}$ is obtained by using (7.5). This estimated covariance matrix can further be used in-stead of the true covariance matrix. If the estimated covariance matrix is the MLE with respect to the available data $\mathbf{y}$, the corresponding test statistic is then called the generalized likelihood ratio.

7.4 Estimation of the complex amplitude of the reflection coefficient $\boldsymbol{\alpha}$

Since only one measurement $\mathbf{y}$ is available to perform the estimation, $\boldsymbol{\alpha}$ is explicitly considered as a deterministic unknown which has to be estimated and this estimate is noted $\hat{\boldsymbol{\alpha}}$. Several methods to perform an estimation of the complex amplitude $\boldsymbol{\alpha}$ of the signal reflected by each clutter patch are described below. The relationships between the estimation methods examined in this section are analyzed in Section 7.6 which gives a way to obtain an estimate of $\mathbf{r}_\alpha$ from $\hat{\boldsymbol{\alpha}}$. Section 7.7.4 discusses the relative performance of these methods.

7.4.1 Least squares estimation

A least squares solution, proposed in [75, 112], consists in finding the $\hat{\boldsymbol{\alpha}}$ which satisfies

$$\min_{\boldsymbol{\alpha}} \|\mathbf{y} - \mathbf{S}\boldsymbol{\alpha}\|^2 \tag{7.13}$$

where $\|\mathbf{a}\|$ denotes the l^2 norm of vector $\mathbf{a}$. If the noise is Gaussian, which is the case here, this approach provides with the MLE [81] and yields the well-known solution [11, 81]

$$\hat{\boldsymbol{\alpha}} = (\mathbf{S}^\dagger\mathbf{S})^{-1}\mathbf{S}^\dagger\mathbf{y}. \tag{7.14}$$

A least squares solution is only feasible if the problem to solve is overdetermined. However, if the number K of clutter patches of interest is larger than the length of the data vector $\mathbf{y}$, the problem becomes underdetermined. It means that there is an upper limit on the number of clutter patches that can be considered in the clutter signal model and hence, significant model errors can occur due to a too low number K of clutter patches.

Even if the problem is overdetermined, a solution only exists if $\mathbf{S}^\dagger\mathbf{S}$ is full rank. In other words, the number of clutter patches K considered in the model (7.3) must be such that $\mathbf{S}^\dagger\mathbf{S}$ is full rank. This approach thus requires an ad-hoc method to estimate the number K of terms to consider, which typically implies a very low number of clutter patches. Hence, the physical meaning of (7.3) is jeopardized and significant model error can occur due to the low value of K. A trivial example of a low rank $\mathbf{S}^\dagger\mathbf{S}$ occurs if the antenna radiation diagram is such that some clutter patches do not actually contribute to $\mathbf{y}$. Obviously, in this case, the corresponding complex amplitude of the reflection coefficients α_i cannot be estimated from the data $\mathbf{y}$. There is yet a more subtle issue: as the spatial resolution of the radar determines the spatial resolution that can be achieved in estimating $\boldsymbol{\alpha}$, particular spatial patterns such as rapidly varying values of α_i along the isorange are attenuated by $\mathbf{S}$ and hardly contribute to $\mathbf{y}$, to such an extent that these spatial patterns cannot be extracted from the noise. These issues are typical for inverse problems [166] and the general method to handle them consists in performing a regularization. Section 7.4.3 discusses a possible solution.

7.4.2 Fast implementation

Implementing (7.14) is computer-intensive and, as noted in [112], a suboptimal approximation to (7.14) is given by[1]

$$\hat{\boldsymbol{\alpha}} = \mathbf{D}^{-1}\mathbf{S}^\dagger\mathbf{y} \tag{7.15}$$

where $\mathbf{D} = \mathrm{tr}(\mathbf{S}^\dagger\mathbf{S})/(NM)$ is a normalization factor. This expression is of course less computer-intensive as no matrix inversion is required. This approximation bears similitude to the method proposed in [94] which considers a ULA and implements $\mathbf{S}$ as several discrete Fourier transforms.

The value of the normalization factor is crucial in obtaining bias-free results. Although the particular value proposed above makes sense, another possible normalization factor, which follows from the developments made in the next sections, is discussed in Section 7.6.

Section 7.6 shows that (7.15) can be seen as a rough approximation of one step of the maximum likelihood estimation of $\mathbf{r}_\alpha$. This estimate is also used in other fields, especially in synthetic aperture radar (SAR) to perform both the azimuth and the range compression [33, Chapter 4] and [45]. Let us call this estimator the crude matched filter (CMF).

7.4.3 Maximum *a posteriori* estimation of $\boldsymbol{\alpha}$

The estimation methods described in the previous sections do not take into account any *a priori* information about $\boldsymbol{\alpha}$. Maximum *a posteriori* (MAP) estimation considers $\boldsymbol{\alpha}$ as a random variable, and as shown shortly hereafter, the MAP estimation permits the introduction of *a priori* knowledge about $\boldsymbol{\alpha}$ in the form of the PDF $p(\boldsymbol{\alpha})$. This *a priori* knowledge is precisely the

[1]The apodization window present in the original paper [112] is dropped in order to ease further discussion although the presence of the window is clearly necessary to avoid spectral leakage.

difference between the MAP estimation and a maximum likelihood (or least squares) approach [81, 169].

In this approach [121], $\hat{\boldsymbol{\alpha}}$ is chosen to maximize the posterior PDF or

$$\hat{\boldsymbol{\alpha}} = \arg\max_{\boldsymbol{\alpha}} p(\boldsymbol{\alpha}|\mathbf{y}). \tag{7.16}$$

by using the Bayes identity

$$p(\boldsymbol{\alpha}|\mathbf{y}) = \frac{p(\mathbf{y}|\boldsymbol{\alpha})p(\boldsymbol{\alpha})}{p(\mathbf{y})}. \tag{7.17}$$

which explicitly takes the *a priori* information $p(\boldsymbol{\alpha})$ about $\boldsymbol{\alpha}$ into account.

As the *a priori* probability $p(\mathbf{y})$ does not depend on $\boldsymbol{\alpha}$, it does not affect the maximum and does not need to be computed. From (7.4), one has $\mathbf{n} = \mathbf{y} - \mathbf{S}\boldsymbol{\alpha}$. As the noise n is assumed to be Gaussian,

$$p(\mathbf{n}) = \frac{1}{\pi^{NM}|\mathbf{R}_n|} e^{-\mathbf{n}^\dagger \mathbf{R}_n^{-1}\mathbf{n}}. \tag{7.18}$$

Hence, (7.4) and (7.18) yield

$$p(\mathbf{y}|\boldsymbol{\alpha}) = \frac{1}{\pi^{NM}|\mathbf{R}_n|} e^{-(\mathbf{y}-\mathbf{S}\boldsymbol{\alpha})^\dagger \mathbf{R}_n^{-1}(\mathbf{y}-\mathbf{S}\boldsymbol{\alpha})}. \tag{7.19}$$

The prior probability $p(\boldsymbol{\alpha})$ expresses the *a priori* knowledge about $\boldsymbol{\alpha}$. As $\boldsymbol{\alpha}$ is the complex amplitude of the signal scattered by the clutter, let us assume that $\boldsymbol{\alpha}$ is independent and complex Gaussian distributed [132], and hence

$$p(\boldsymbol{\alpha}) = \frac{1}{\pi^{K}|\tilde{\mathbf{R}}_\alpha|} e^{-\boldsymbol{\alpha}^\dagger \tilde{\mathbf{R}}_\alpha^{-1}\boldsymbol{\alpha}} \tag{7.20}$$

where $\tilde{\mathbf{R}}_\alpha$, expressing the *a priori* knowledge about $\boldsymbol{\alpha}$, is taken diagonal.

Finally, combining (7.17), (7.19) and (7.20) yields

$$p(\boldsymbol{\alpha}|\mathbf{y}) \propto \frac{1}{\pi^{N}M|\mathbf{R}_n|\pi^{K}|\tilde{\mathbf{R}}_\alpha|} e^{-(\mathbf{y}-\mathbf{S}\boldsymbol{\alpha})^\dagger \mathbf{R}_n^{-1}(\mathbf{y}-\mathbf{S}\boldsymbol{\alpha})-\boldsymbol{\alpha}^\dagger \tilde{\mathbf{R}}_\alpha^{-1}\boldsymbol{\alpha}}. \tag{7.21}$$

As the matrices $\mathbf{R}_n$ and $\tilde{\mathbf{R}}_\alpha$ are positive definite, the maximum is reached for

$$\hat{\boldsymbol{\alpha}} = (\mathbf{S}^\dagger \mathbf{R}_n^{-1}\mathbf{S} + \tilde{\mathbf{R}}_\alpha^{-1})^{-1}\mathbf{S}^\dagger \mathbf{R}_n^{-1}\mathbf{y}. \tag{7.22}$$

Note that if no *a priori* information is available, i.e., $\tilde{\mathbf{R}}_\alpha^{-1} = \mathbf{0}$, and if the noise variance is identical for all sensors and pulses, this estimate is the same as (7.14).

Formally, if $\mathbf{R}_n$ and $\tilde{\mathbf{R}}_\alpha$ are proportional to the identity matrix, the MAP estimate can be seen as a regularized form of the least squares solution (7.14) where $\boldsymbol{\alpha}$ with small magnitudes are favored [6]. The term $\tilde{\mathbf{R}}_\alpha^{-1}$ thus acts as a regularization term by constraining the magnitude of α_i if $\mathbf{S}^\dagger \mathbf{R}_n^{-1}\mathbf{S}$ is rank deficient.

More generally speaking, the regularization term favors solutions $\hat{\alpha}_i$ with small magnitudes. In this sense, the solution with the smallest magnitude is selected whenever several solutions are possible.

7.5 Estimation of the scattering coefficient $\mathbf{r}_\alpha$

This section establishes the equation of the maximum likelihood estimate of $\mathbf{r}_\alpha$ (or equivalently of $\mathbf{R}_\alpha$), called the trace equation, and discusses practical methods to solve it.

7.5.1 Invariance principle of the likelihood

Here, the intention consists in estimating $\mathbf{R}$ from the data $\mathbf{y}$ by taking into account the structure of $\mathbf{R}$ given in (7.5). Clearly, $\mathbf{R}$ depends on $\mathbf{R}_\alpha$ and is therefore denoted by $\mathbf{R}(\mathbf{R}_\alpha)$.

Let us make the estimation by computing the maximum likelihood estimate of $\mathbf{R}$. It however proves to be easier to compute the maximum likelihood $\hat{\mathbf{R}}_\alpha$ of $\mathbf{R}_\alpha$. Intuitively, $\hat{\mathbf{R}}_\alpha$ would then be used in (7.5) and it is expected to be yield maximum likelihood estimate of $\mathbf{R}$. This is indeed known as the invariance principle of the likelihood [81, 136] which states:

> "The maximum likelihood estimate of $\mathbf{R} = \mathbf{R}(\mathbf{R}_\alpha)$ is given by
>
> $$\hat{\mathbf{R}} = \mathbf{R}(\hat{\mathbf{R}}_\alpha) \tag{7.23}$$
>
> where $\hat{\mathbf{R}}_\alpha$ is the maximum likelihood estimate of $\mathbf{R}_\alpha$."

The above formulation requires that there exists an injective mapping between $\mathbf{R}_\alpha$ and $\mathbf{R}$. If this is not verified, then $\hat{\mathbf{R}}$ maximizes the modified likelihood function defined as

$$L'(\mathbf{R}) = \max_{\{\mathbf{R}_\alpha : \mathbf{R} = \mathbf{R}(\mathbf{R}_\alpha)\}} L(\mathbf{R}_\alpha). \tag{7.24}$$

Here, (7.5) is not invertible and there might thus exist different values of $\mathbf{R}_\alpha$ giving the same $\mathbf{R}$. As explained in the next section, the likelihood of $\mathbf{R}_\alpha$ depends on $\mathbf{R}_\alpha$ through $\mathbf{R}$. Hence, the different values of $\mathbf{R}_\alpha$ giving rise to the same $\mathbf{R}$ have the same likelihood and the modified likelihood function (7.24) is then equal to the usual likelihood function.

7.5.2 Maximum likelihood estimation – Trace equation

The equation that the maximum likelihood estimate must satisfy is now developed. It will be solved in the two next sections.

The estimate of $\mathbf{R}_\alpha$, which is the most compatible with the measurements $\mathbf{y}$, is obtained by maximizing the probability $p(\mathbf{R}_\alpha|\mathbf{y})$ hence implicitly considering $\mathbf{R}_\alpha$ as a random variable. The Bayes identity gives

$$p(\mathbf{R}_\alpha|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{R}_\alpha)p(\mathbf{R}_\alpha)}{p(\mathbf{y})}, \tag{7.25}$$

where $p(\mathbf{R}_\alpha)$ is some *a priori* probability of $\mathbf{R}_\alpha$, and $p(\mathbf{y})$ is the *a priori* probability of the measurement $\mathbf{y}$, independent of $\mathbf{R}_\alpha$. If a flat *a priori* probability for $\mathbf{R}_\alpha$ is used, the optimum estimate of $\mathbf{R}_\alpha$ can be found by maximizing $L(\mathbf{R}_\alpha) = p(\mathbf{y}|\mathbf{R}_\alpha)$ which is classically called the likelihood of $\mathbf{R}_\alpha$.

If $\mathbf{y}$ is Gaussian with covariance matrix $\mathbf{R}$ given by (7.5),

$$L(\mathbf{R}_\alpha) = p(\mathbf{y}|\mathbf{R}_\alpha) = \frac{1}{\pi^{NM}|\mathbf{R}|} e^{-\mathbf{y}^\dagger \mathbf{R}^{-1}\mathbf{y}}, \tag{7.26}$$

where $|\mathbf{R}|$ denotes the determinant of the matrix $\mathbf{R}$. As (see [25]) $\mathbf{y}^\dagger\mathbf{R}^{-1}\mathbf{y} = \mathrm{tr}(\mathbf{y}^\dagger\mathbf{R}^{-1}\mathbf{y}) = \mathrm{tr}(\mathbf{R}^{-1}\mathbf{y}\mathbf{y}^\dagger)$, the logarithm of the likelihood $l(\mathbf{R}_\alpha)$ is given by

$$l(\mathbf{R}_\alpha) = \ln p(\mathbf{y}|\mathbf{R}_\alpha) = -\ln|\mathbf{R}| - \mathrm{tr}(\mathbf{R}^{-1}\mathbf{y}\mathbf{y}^\dagger) + \kappa. \tag{7.27}$$

where κ is a constant, independent of $\mathbf{R}_\alpha$.

A necessary condition so that $l(\mathbf{R}_\alpha)$ is maximum is

$$\frac{\partial l(\mathbf{R}_\alpha)}{\partial r_{\alpha_i}} = 0 \tag{7.28}$$

for all i. Following [25, 156], if $\mathbf{B}$ is a regular square matrix and b a scalar, then

$$\frac{\partial \ln|\mathbf{B}|}{\partial b} = \mathrm{tr}\left(\mathbf{B}^{-1}\frac{\partial \mathbf{B}}{\partial b}\right) \tag{7.29}$$

hence,

$$\frac{\partial \ln|\mathbf{R}|}{\partial r_{\alpha_i}} = \mathrm{tr}\left(\mathbf{R}^{-1}\frac{\partial \mathbf{R}}{\partial r_{\alpha_i}}\right) = \mathrm{tr}(\mathbf{R}^{-1}\mathbf{s}_i\mathbf{s}_i^\dagger) = \mathrm{tr}(\mathbf{s}_i^\dagger\mathbf{R}^{-1}\mathbf{s}_i). \tag{7.30}$$

Expressing it for all i, with $\{\mathbf{B}\}_d$ denoting the column vector consisting of the diagonal elements of $\mathbf{B}$, yields

$$\frac{\partial \ln|\mathbf{R}|}{\partial \mathbf{r}_\alpha} = \{\mathbf{S}^\dagger\mathbf{R}^{-1}\mathbf{S}\}_d. \tag{7.31}$$

Using the property (see [25])

$$\frac{\partial \mathbf{B}^{-1}}{\partial b} = -\mathbf{B}^{-1}\frac{\partial \mathbf{B}}{\partial b}\mathbf{B}^{-1} \tag{7.32}$$

yields

$$\frac{\partial\, \mathrm{tr}(\mathbf{R}^{-1}\mathbf{y}\mathbf{y}^\dagger)}{\partial r_{\alpha_i}} = \mathrm{tr}\left(\frac{\partial \mathbf{R}^{-1}}{\partial r_{\alpha_i}}\mathbf{y}\mathbf{y}^\dagger\right) = \mathrm{tr}(-\mathbf{R}^{-1}\mathbf{s}_i\mathbf{s}_i^\dagger\mathbf{R}^{-1}\mathbf{y}\mathbf{y}^\dagger) = -\,\mathrm{tr}(\mathbf{s}_i^\dagger\mathbf{R}^{-1}\mathbf{y}\mathbf{y}^\dagger\mathbf{R}^{-1}\mathbf{s}_i) \tag{7.33}$$

or by expressing it for all i

$$\frac{\partial\, \mathrm{tr}(\mathbf{R}^{-1}\mathbf{y}\mathbf{y}^\dagger)}{\partial \mathbf{r}_\alpha} = -\{\mathbf{S}^\dagger\mathbf{R}^{-1}\mathbf{y}\mathbf{y}^\dagger\mathbf{R}^{-1}\mathbf{S}\}_d. \tag{7.34}$$

The derivative of (7.27) finally becomes

$$\frac{\partial l(\mathbf{R}_\alpha)}{\partial \mathbf{r}_\alpha} = \{\mathbf{S}^\dagger\mathbf{R}^{-1}(\mathbf{y}\mathbf{y}^\dagger - \mathbf{R})\mathbf{R}^{-1}\mathbf{S}\}_d. \tag{7.35}$$

Equation (7.28) can thus be rewritten for all i in vector form as

$$\{\mathbf{S}^\dagger\mathbf{R}^{-1}(\mathbf{y}\mathbf{y}^\dagger - \mathbf{R})\mathbf{R}^{-1}\mathbf{S}\}_d = 0, \tag{7.36}$$

which is a particular case of the trace equation [159]. This equation must be solved for $\mathbf{R}_\alpha$. It is nonlinear in $\mathbf{R}_\alpha$ and no closed-form solution is known. Although the classical way to obtain the maximum likelihood estimate is the Expectation-Maximization algorithm [37] presented in the next section, a more efficient iterative method is presented in Section 7.5.4.

7.5.3 Expectation-Maximization

The Expectation-Maximization (EM) algorithm [37] can be used to find the value of $\mathbf{R}_\alpha$ that maximizes (7.27) [47, 91]. In the EM algorithm, the coefficients $\boldsymbol{\alpha}$ are called the complete data while $\mathbf{y}$ are the incomplete data and (7.4) provides a mapping between them.

The EM algorithm is an iterative algorithm consisting in a succession of expectation (E) and maximization (M) steps. In the E-step, the expectation of the log-likelihood of the complete data $l_{cd}(\mathbf{R}_\alpha|\boldsymbol{\alpha})$ conditioned upon an estimate $\mathbf{R}_\alpha^{(k)}$ of $\mathbf{R}_\alpha$, and upon the incomplete data $\mathbf{y}$ is computed,

$$\Xi(\mathbf{R}_\alpha|\mathbf{R}_\alpha^{(k)}) = E\{l_{cd}(\mathbf{R}_\alpha|\boldsymbol{\alpha})|\mathbf{R}_\alpha^{(k)}, \mathbf{y}\}. \tag{7.37}$$

In the M-step, the value of $\mathbf{R}_\alpha$ that maximizes $\Xi(\mathbf{R}_\alpha|\mathbf{R}_\alpha^{(k)})$ is computed

$$\mathbf{R}_\alpha^{(k+1)} = \underset{\mathbf{R}_\alpha}{\operatorname{argmax}}\, \Xi(\mathbf{R}_\alpha|\mathbf{R}_\alpha^{(k)}). \tag{7.38}$$

These two steps are repeated until convergence.

As $\boldsymbol{\alpha}$ is complex Gaussian distributed with covariance matrix $\mathbf{R}_\alpha$, according to (7.20), a development analogous to (7.27) yields

$$l_{cd}(\mathbf{R}_\alpha|\boldsymbol{\alpha}) = -\ln|\mathbf{R}_\alpha| - \operatorname{tr}(\mathbf{R}_\alpha^{-1}\boldsymbol{\alpha}\boldsymbol{\alpha}^\dagger) + \kappa' \tag{7.39}$$

where κ' is a known constant independent of $\mathbf{R}_\alpha$ and (7.37) can be evaluated as

$$\Xi(\mathbf{R}_\alpha|\mathbf{R}_\alpha^{(k)}) = -\ln|\mathbf{R}_\alpha| - \operatorname{tr}(\mathbf{R}_\alpha^{-1}E\{\boldsymbol{\alpha}\boldsymbol{\alpha}^\dagger|\mathbf{R}_\alpha^{(k)}, \mathbf{y}\}) + \kappa'. \tag{7.40}$$

As the term $E\{\boldsymbol{\alpha}\boldsymbol{\alpha}^\dagger|\mathbf{R}_\alpha^{(k)}, \mathbf{y}\}$ does not depend on $\mathbf{R}_\alpha$, taking the derivative of $\Xi(\mathbf{R}_\alpha|\mathbf{R}_\alpha^{(k)})$ with respect to $\mathbf{r}_\alpha$, and using (7.29) and (7.32), yield

$$\frac{\partial\Xi(\mathbf{R}_\alpha|\mathbf{R}_\alpha^{(k)})}{\partial\mathbf{r}_\alpha} = -\{\mathbf{R}_\alpha^{-1}\}_d + \{\mathbf{R}_\alpha^{-1}\mathbf{R}_\alpha^{-1}E\{\boldsymbol{\alpha}\boldsymbol{\alpha}^\dagger|\mathbf{R}_\alpha^{(k)}, \mathbf{y}\}\}_d. \tag{7.41}$$

Setting this derivative equal to zero to find the maximum of $\Xi(\mathbf{R}_\alpha|\mathbf{R}_\alpha^{(k)})$ yields

$$\mathbf{r}_\alpha^{(k+1)} = \{E\{\boldsymbol{\alpha}\boldsymbol{\alpha}^\dagger|\mathbf{R}_\alpha^{(k)}, \mathbf{y}\}\}_d, \tag{7.42}$$

which is a set of K equations. Equation i of this set is

$$r_{\alpha_i}^{(k+1)} = E\{|\alpha_i|^2|\mathbf{R}_\alpha^{(k)}, \mathbf{y}\}. \tag{7.43}$$

The term $E\{\boldsymbol{\alpha}\boldsymbol{\alpha}^\dagger|\mathbf{R}_\alpha^{(k)}, \mathbf{y}\}$ appearing in (7.42) can be expressed as a function of the conditional mean $\bar{\boldsymbol{\alpha}} = E\{\boldsymbol{\alpha}|\mathbf{R}_\alpha^{(k)}, \mathbf{y}\}$ and the conditional covariance of $\boldsymbol{\alpha}$

$$E\{\boldsymbol{\alpha}\boldsymbol{\alpha}^\dagger|\mathbf{R}_\alpha^{(k)}, \mathbf{y}\} = \operatorname{diag}[\{\bar{\boldsymbol{\alpha}}\bar{\boldsymbol{\alpha}}^\dagger\}_d + \{\operatorname{cov}\{\boldsymbol{\alpha}|\mathbf{R}_\alpha^{(k)}, \mathbf{y}\}\}_d] \tag{7.44}$$

where the conditional mean and the conditional covariance are well-known results from estimation theory [11, 47, 81], respectively given by

$$\bar{\boldsymbol{\alpha}} = E\{\boldsymbol{\alpha}|\mathbf{R}_\alpha^{(k)}, \mathbf{y}\} \tag{7.45}$$

$$= \mathbf{R}_\alpha^{(k)}\mathbf{S}^\dagger\mathbf{R}^{-1}\mathbf{y} \tag{7.46}$$

with $\mathbf{R} = \mathbf{R}(\mathbf{R}_\alpha)|_{\mathbf{R}_\alpha^{(k)}}$, and by

$$\mathrm{cov}\{\boldsymbol{\alpha}|\mathbf{R}_\alpha^{(k)}, \mathbf{y}\} = \mathbf{R}_\alpha^{(k)} - \mathbf{R}_\alpha^{(k)}\mathbf{S}^\dagger\mathbf{R}^{-1}\mathbf{S}\mathbf{R}_\alpha^{(k)}. \tag{7.47}$$

Hence, (7.42) can be rewritten as

$$\mathbf{r}_\alpha^{(k+1)} = \{\mathbf{R}_\alpha^{(k)}\mathbf{S}^\dagger\mathbf{R}^{-1}\mathbf{y}\mathbf{y}^\dagger\mathbf{R}^{-1}\mathbf{S}\mathbf{R}_\alpha^{(k)}\}_d + \{\mathbf{R}_\alpha^{(k)}\}_d - \{\mathbf{R}_\alpha^{(k)}\mathbf{S}^\dagger\mathbf{R}^{-1}\mathbf{S}\mathbf{R}_\alpha^{(k)}\}_d \tag{7.48}$$

where the first term is recognized as the MAP estimate of $\boldsymbol{\alpha}$ and the two last terms are the covariance of $\boldsymbol{\alpha}$.

Using the expression of $\frac{\partial l(\mathbf{R}_\alpha)}{\partial \mathbf{r}_\alpha}$ obtained in the previous section, finally yields one iteration of the EM algorithm given by

$$\mathbf{r}_\alpha^{(k+1)} = \mathbf{r}_\alpha^{(k)} + \mathbf{r}_\alpha^{(k)} \circ \left.\frac{\partial l(\mathbf{R}_\alpha)}{\partial \mathbf{r}_\alpha}\right|_{\mathbf{R}_\alpha^{(k)}} \circ \mathbf{r}_\alpha^{(k)}. \tag{7.49}$$

where $\circ$ denotes the Hadamard element-wise product. This latter expression shows that, in this application, the EM algorithm can be interpreted as a weighted gradient descent method to find the value of $\mathbf{R}_\alpha$ that maximizes the log-likelihood $l(\mathbf{R}_\alpha)$.

7.5.4 Iterative solution

An iterative resolution of (7.36) is proposed in [156] and is now adapted to this application. As illustrated in Section 7.7.3, this method converges faster than the EM method.

Noting

$$\mathbf{F} = \mathbf{R}_\alpha\mathbf{S}^\dagger\mathbf{R}^{-1}, \tag{7.50}$$

and by decomposing $\mathbf{R}$ according to (7.5), (7.35) can be rewritten as

$$\mathbf{r}_\alpha \circ \frac{\partial l(\mathbf{R}_\alpha)}{\partial \mathbf{r}_\alpha} \circ \mathbf{r}_\alpha = \{\mathbf{F}\mathbf{y}\mathbf{y}^\dagger\mathbf{F}^\dagger\}_d - \{\mathbf{F}\mathbf{S}\mathbf{R}_\alpha\mathbf{S}^\dagger\mathbf{F}^\dagger\}_d - \{\mathbf{F}\mathbf{R}_n\mathbf{F}^\dagger\}_d \tag{7.51}$$

$$= \{\mathbf{F}\mathbf{y}\mathbf{y}^\dagger\mathbf{F}^\dagger\}_d - \mathbf{T}\mathbf{R}_\alpha - \{\mathbf{F}\mathbf{R}_n\mathbf{F}^\dagger\}_d. \tag{7.52}$$

The first term of this last expression equals $\{\bar{\boldsymbol{\alpha}}\bar{\boldsymbol{\alpha}}^\dagger\}_d$ and $\mathbf{T}$ is defined [123] similarly[2] as in [156],

$$\mathbf{T} = \mathrm{diag}\{\{\mathbf{F}\mathbf{S}\mathbf{R}_\alpha\mathbf{S}^\dagger\mathbf{F}^\dagger\}_d\}\mathbf{R}_\alpha^{-1}. \tag{7.53}$$

[2]This definition of $\mathbf{T}$ slightly differs from the definition given in [156], where $\mathbf{T} = \mathrm{diag}\{\{\mathbf{F}\mathbf{S}\mathbf{S}^\dagger\mathbf{F}^\dagger\}_d\}$ which makes the introduction of $\mathbf{T}$ more natural. However, the latter expression results from a mistake in [156].

In order to give an interpretation to $\mathbf{T}$, let us consider the case for which $\mathbf{R}_\alpha$ is proportional to the identity matrix, $\mathbf{R}_\alpha = \epsilon\mathbf{I}$. In this case, $\mathbf{T}$ reduces to

$$\mathbf{T} = \mathrm{diag}\{\{\mathbf{F}\mathbf{S}\mathbf{S}^\dagger\mathbf{F}^\dagger\}_d\} \tag{7.54}$$

which shows that $\mathbf{T}$ can be seen as a normalization factor. In particular, $\mathbf{T}$ is, for large clutter to noise ratio, invariant to a scale factor on $\mathbf{R}_\alpha$.

Setting (7.52) equal to zero yields

$$\mathbf{r}_\alpha = \mathbf{T}^{-1}(\{\mathbf{F}\mathbf{y}\mathbf{y}^\dagger\mathbf{F}^\dagger\}_d - \{\mathbf{F}\mathbf{R}_n\mathbf{F}^\dagger\}_d). \tag{7.55}$$

Consequently, the following iterative solution is proposed [156]

$$\mathbf{r}_\alpha^{(k+1)} = \mathbf{r}_\alpha^{(k)} + \tau \left[\mathbf{T}^{-1}(\{\mathbf{F}\mathbf{y}\mathbf{y}^\dagger\mathbf{F}^\dagger\}_d - \{\mathbf{F}\mathbf{R}_n\mathbf{F}^\dagger\}_d) - \mathbf{r}_\alpha\right]\Big|_{\mathbf{r}_\alpha^{(k)}}. \tag{7.56}$$

By taking (7.52) into account, this last equation can be rewritten as

$$\mathbf{r}_\alpha^{(k+1)} = \mathbf{r}_\alpha^{(k)} + \tau \left[\mathbf{T}^{-1}\left(\mathbf{r}_\alpha \circ \frac{\partial l(\mathbf{R}_\alpha)}{\partial \mathbf{r}_\alpha} \circ \mathbf{r}_\alpha\right)\right]\Bigg|_{\mathbf{r}_\alpha^{(k)}}. \tag{7.57}$$

which shows that the iterative scheme proposed in [156] is equivalent to a modified gradient descent where the gradient is multiplied by a matrix. Moreover, taking $\tau = 1$ in (7.57) results in

$$\mathbf{r}_\alpha^{(k+1)} = \left[\mathbf{T}^{-1}(\{\mathbf{F}\mathbf{y}\mathbf{y}^\dagger\mathbf{F}^\dagger\}_d - \{\mathbf{F}\mathbf{R}_n\mathbf{F}^\dagger\}_d)\right]\big|_{\mathbf{r}_\alpha^{(k)}}. \tag{7.58}$$

The fixed points of this equation satisfy (7.55), which might be interpreted as an iterative solution to (7.55). Note that while convergence is achieved in practice, the convergence of (7.58) is not claimed.

7.5.5 Regularization

In the considered application, $\mathbf{S}\mathbf{R}_\alpha\mathbf{S}^\dagger$ is not full rank, which implies that there is no unique solution for $\mathbf{R}_\alpha$. This kind of problem is "ill conditioned" and regularization is necessary to obtain a solution [166]. Regularization is obtained by introducing *a priori* knowledge about the solution $\mathbf{R}_\alpha$. This can be done, e.g., by imposing directly an *a priori* probability density of $\mathbf{R}_\alpha$ in the likelihood to minimize [91, 156, 157] or by restricting the space in which the solution is to be found, e.g., by decomposing the covariance matrix $\mathbf{R}_\alpha$ in a basis in order to impose its structure [25, 117, 116]. The latter decomposition is particularly suited for cases where *a priori* knowledge about the ground cover is known; the covariance matrix can then be decomposed in a set of matrices that takes into account this knowledge. It is also worth mentioning that a review of various regularization methods applied to the EM algorithm can be found in [90, 91].

A particular regularization is one that consists in adding a spatial smoothing step after the computation of the M-step of the EM algorithm. [91, 156] shows that, in some particular cases, there is an equivalence between the imposition of some prior probability density $p(\mathbf{R}_\alpha)$ and the addition of a spatial smoothing step. This motivates the introduction of a spatial smoothing step in (7.58):

$$\mathbf{r}_\alpha^{(k+1)} = \mathbf{W}\left[\mathbf{T}^{-1}(\{\mathbf{F}\mathbf{y}\mathbf{y}^\dagger\mathbf{F}^\dagger\}_d - \{\mathbf{F}\mathbf{R}_n\mathbf{F}^\dagger\}_d)\right]|_{R_\alpha^{(k)}} \tag{7.59}$$

and in (7.49):

$$\mathbf{r}_\alpha^{(k+1)} = \mathbf{W}\left(\mathbf{r}_\alpha^{(k)} + \mathbf{r}_\alpha^{(k)} \circ \left.\frac{\partial l(\mathbf{R}_\alpha)}{\partial \mathbf{r}_\alpha}\right|_{\mathbf{R}_\alpha^{(k)}} \circ \mathbf{r}_\alpha^{(k)}\right), \tag{7.60}$$

where the square matrix $\mathbf{W}$ is a circulant smoothing matrix. The matrix $\mathbf{W}$ is circulant because the scattering coefficients r_{α_i} are located along a closed isorange on the ground. The effect of $\mathbf{W}$ is to perform a weighted averaging of r_{α_i} with the neighboring scattering coefficients.

Note that although the spatial averaging is formally motivated by the need to include *a priori* knowledge about $\mathbf{R}_\alpha$, it is also reasonable from a physical point of view. Indeed, the spatial resolution of the radar limits the observable spatial variations of r_{α_i} and it is thus reasonable to remove the fast variation of $\mathbf{r}_\alpha$ by spatial averaging. A similar situation arises for instance in synthetic aperture radar (SAR) where the scattering coefficients need also to be estimated. The resulting variance is called "speckle" and the usual way of reducing it consists in performing "multilooking" or spatial averaging [132]. The classical spatially-adaptive multilooking methods developed in that field can be applied here. For instance, in order to preserve the structure possibly present in the clutter reflectivity map and due to the ground cover inhomogeneity, more sophisticated spatial filtering methods such as [97] could be used.

7.5.6 Discussion

7.5.6.1 Application to scattering coefficient maps estimation

Equations (7.3) and (7.5) can be written for each range-gate. Note that for bistatic scenarios, $\mathbf{S}$ typically depends on the range. If the measurements $\mathbf{y}$ at different ranges are uncorrelated, i.e., if the range sidelobes are negligible, the estimation of the unregularized $\mathbf{R}_\alpha$ is decoupled in range, which simplifies the computations.

However, the physics clearly imposes a spatial constraint on the scattering coefficients r_{α_i} in the sense that a smooth evolution of the scattering coefficients should be favored, both in azimuth (cross-range) and in range. Hence, the smoothing step discussed in the previous section needs to be extended to smooth the estimate of $r_{\alpha_i}^{(k)}$ also across different ranges.

7.5.6.2 Fast implementation

Equation (7.59), although providing a solution to the problem, results in an extremely computer-intensive algorithm. It turns out that one single iteration of (7.59) already yields a very acceptable estimate of $\mathbf{r}_\alpha$, as shown in Section 7.7.3. This is partly due to the correcting factor $\mathbf{T}$ that hardly depends on $\mathbf{R}_\alpha$. Therefore, let us consider a single iteration of (7.59) as estimator.

Moreover, if the clutter-to-noise ratio is high enough, the contribution of the second term in (7.59) may be neglected. At one range, an estimate of $\mathbf{R}_\alpha$ is thus obtained by computing

$$\hat{\mathbf{r}}_\alpha = \mathbf{W}\left(\mathbf{T}^{-1}\{\bar{\boldsymbol{\alpha}}\bar{\boldsymbol{\alpha}}^\dagger\}_d\right)\Big|_{\mathbf{R}_\alpha^{(0)}}, \tag{7.61}$$

where $\bar{\boldsymbol{\alpha}}$ is the MAP estimate of $\boldsymbol{\alpha}$ conditioned upon $\mathbf{R}_\alpha^{(0)}$, i.e., where $\tilde{\mathbf{R}}_\alpha = \mathbf{R}_\alpha^{(0)}$.

7.5.6.3 Final algorithm

The proposed algorithm thus consists in two steps. The first step estimates $\bar{\boldsymbol{\alpha}}$ at each range and computes $\bar{\mathbf{r}}_\alpha = T^{-1}\{\bar{\boldsymbol{\alpha}}\bar{\boldsymbol{\alpha}}^\dagger\}_d$ at each range independently. The second step performs a local spatial averaging of the estimated scattering coefficients.

7.6 Relationship between the estimation algorithms

The relationship between the different estimation algorithms presented in the previous sections is now discussed.

7.6.1 Summary of the different algorithms

In the previous sections, four estimation methods were presented:

The Expectation-Maximization estimator of the MLE of R_α (EM), used as a reference is however the most computer-intensive as it is iterative.

A single iteration of the MLE of R_α (MLE) which is another way of computing the MLE of $\mathbf{R}_\alpha$.

The MAP estimate of $\boldsymbol{\alpha}$ (MAP) described by (7.22)

$$\hat{\mathbf{r}}_\alpha = \mathbf{W}\{\hat{\boldsymbol{\alpha}}\hat{\boldsymbol{\alpha}}^\dagger\}_d \tag{7.62}$$

where $\mathbf{W}$ is a spatial averaging window and

$$\hat{\boldsymbol{\alpha}} = (\mathbf{S}^\dagger\mathbf{R}_n^{-1}\mathbf{S} + \tilde{\mathbf{R}}_\alpha^{-1})^{-1}\mathbf{S}^\dagger\mathbf{R}_n^{-1}\mathbf{y}. \tag{7.63}$$

The Crude Matched Filter (CMF), described by

$$\hat{\mathbf{r}}_\alpha = \mathbf{W}\{\hat{\boldsymbol{\alpha}}\hat{\boldsymbol{\alpha}}^\dagger\}_d \tag{7.64}$$

where $\mathbf{W}$ is a spatial averaging window and (7.15)

$$\hat{\boldsymbol{\alpha}} = \mathbf{D}^{-1}\mathbf{S}^\dagger\mathbf{y}. \tag{7.65}$$

where $\mathbf{D}$ is a diagonal normalization matrix.

It is worth noting that the spatial averaging is extended to the two last methods in order to provide a fair comparison. The motivation of this extension is that $\mathbf{r}_\alpha = \{E\{\boldsymbol{\alpha}\boldsymbol{\alpha}^\dagger\}\}_d$ and that the expectation operator can be approached by performing a spatial averaging, assuming independence of the α and local homogeneity. This procedure is commonly used in SAR [132], where single look complex coefficients are spatially averaged to get an estimate of the scattering coefficients $\hat{\mathbf{r}}_\alpha$.

7.6.2 A single iteration of the MLE of R_α

As shown after the discussion of the performances, the iterative method is such that a single iteration yields acceptable results.

7.6.3 The MAP estimate of $\boldsymbol{\alpha}$

While the obtained estimate $\hat{\boldsymbol{\alpha}}$ is correct, the deduced estimate $\hat{\mathbf{r}}_\alpha$ lacks the covariance term of $\boldsymbol{\alpha}$. More specifically, the estimator $\hat{\mathbf{r}}_\alpha = \mathbf{W}\{\hat{\boldsymbol{\alpha}}\hat{\boldsymbol{\alpha}}^\dagger\}_d$ is biased, even if the true value of the covariance matrix of $\boldsymbol{\alpha}$ is used in the *a priori* knowledge, which leads to an under-estimation of the scattering coefficients, as shown in the next section.

7.6.4 The Crude Matched Filter

The crude matched filter is best motivated as an approximation of the MAP estimator [112]. Another interpretation [94] consists in considering that the coefficients $\mathbf{r}_\alpha$ could be obtained by performing a spectral analysis of $\mathbf{y}$. This is true as long as the vectors $\mathbf{s}_k$ are orthogonal (orthonormality would also solve the problem of determining the normalization factor). Orthogonality is only achieved for particular choices of the Doppler and spatial frequencies for each $\mathbf{s}_i$. And, additionally, these frequencies are required to be located on the clutter power spectrum locus. Generally, it is not possible to fulfill these two requirements simultaneously. This lack of orthogonality introduces cross-product terms and additionally degrades the estimate of the scattering coefficients. Further, the cross-product terms induce an azimuth-dependent bias.

The CMF (7.15)

$$\hat{\alpha} = \mathbf{D}^{-1}\mathbf{S}^\dagger\mathbf{y} \tag{7.66}$$

requires a normalization $\mathbf{D}$. Considering

$$\mathbf{D} = \mathrm{tr}(\mathbf{S}\mathbf{S}^\dagger)/NM \tag{7.67}$$

takes into account the fact that the vectors $\mathbf{s}_i$ are not necessarily normalized as they include the geometric factors of the radar range equation. Another normalization is proposed in [156]

$$\mathbf{D} = \sqrt{\frac{\mathrm{tr}(\mathbf{S}^\dagger\mathbf{S}\mathbf{S}^\dagger\mathbf{S})}{\mathrm{tr}(\mathbf{F}^\dagger\mathbf{S}\mathbf{S}^\dagger\mathbf{F})}}. \tag{7.68}$$

This normalization factor is also sensible and indeed actually better reduces the bias, particularly if the $\mathbf{s}_i$'s do not have equal amplitudes. However, its computation requires the computation of the MAP filter $\mathbf{F}$, and thus as much computations as the computation of the MAP estimate itself, hence voiding the whole point of the CMF. From a practical standpoint it is thus not feasible. Note that neither [112] nor [94] discuss in detail the normalization factor they consider and this normalization factor is certainly not the expression given by (7.68).

7.7 Performance comparison

7.7.1 Influence of the accuracy of the scattering coefficient estimation

This section assesses the influence of a misestimation of the clutter power on the detection performance. Figure 7.3 illustrates the loss in SINR due to a power misestimation of the clutter for a nominal clutter to noise ratio of 20dB in the case of a monostatic scenario with a ULA. The figure presents a cut in the SINR loss at $\nu_s = 0$. The loss is 0 if there is no misestimation. Below

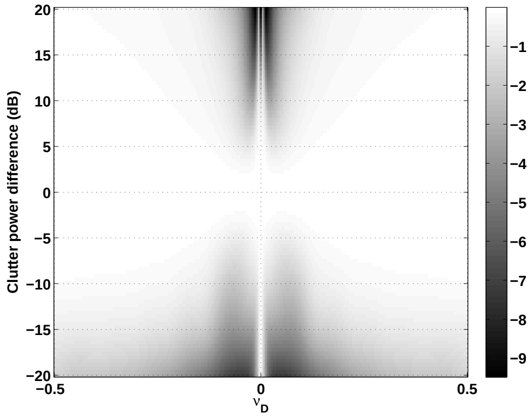


Figure 7.3: Influence of a misestimation of the clutter power on the SINR loss.

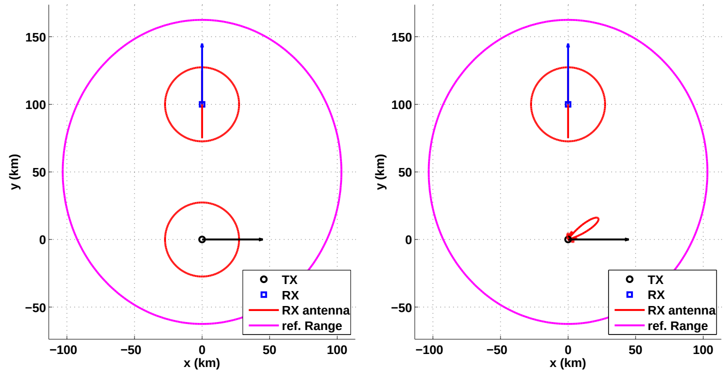
approximately ±5dB of power mismatch between the true clutter power and the assumed clutter power, there is practically no SINR loss. As the SINR loss is always negative, over-nulling always occurs.

7.7.2 Scenario of interest

The scenario of interest is a bistatic setup where the transmitter is located at the origin and the receiver is located at $(0, 100)$. The transmitter platform is flying east while the receiver is flying north. Two cases are considered and illustrated in Figure 7.4. In the first case, omnidirectional antennas are considered while in the second case, a directive Tx antenna is considered in order to assess the performance of the method if only a part of the scattering coefficients can be estimated.

The clutter to noise ratio is 20dB. A homogeneous ground cover with scattering coefficients taken equal to 0dB is first considered in order to assess the performance of the method under the assumption of a homogeneous clutter. Next, a checkerboard ground cover, with scattering coefficients switching between +5dB and -5dB, is considered in order to assess the performance of the method if the homogeneity assumption is not verified.

A separable square Hanning window of length 5 is used for the spatial averaging.



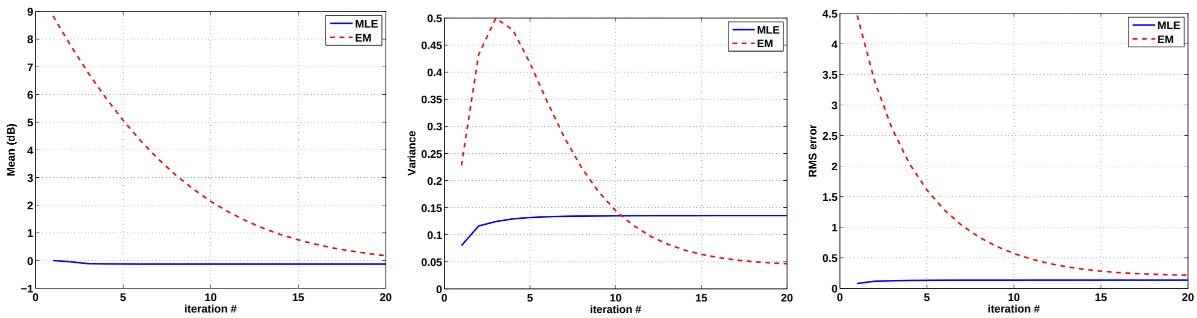
(a) (b)

Figure 7.4: Scenario of interest (a) with omnidirectional antennas and (b) with a directive Tx antenna.

7.7.3 Comparison of the convergence behavior of the iterative algorithms

First, the convergence of the two iterative algorithms to compute the maximum likelihood estimate of $\mathbf{r}_\alpha$ and presented in Section 7.5 is analyzed. A homogeneous ground cover is considered.

The initial value of the covariance matrix $\mathbf{R}_\alpha^{(0)}$ is taken equal to $r\mathbf{I}$ where r has an offset of +10dB with respect to the true average value. An overestimation of the initial value of the covariance matrix means that little *a priori* information about $\mathbf{r}_\alpha$ is introduced. Figure 7.5 presents



(a) (b) (c)

Figure 7.5: Evolution of (a) the mean value, (b) the spatial variance of $\mathbf{r}_\alpha^{(k)}$ and (c) the RMS error w.r.t. the true value as a function of the iteration number.

different metrics measuring the convergence. Figure 7.5 (a) presents the mean value of the estimated scattering coefficient, i.e., the value which would be obtained with a large spatial filter and which is used to assess the bias of the estimate, since the true value is spatially constant. Figure 7.5 (b) represents the variance (around the mean) which can be used to assess the spatial homogeneity of the estimate. Figure 7.5 (c) gives the usual RMS value. As the RMS value combines the bias and the variance around the mean, it is less informative in the case of a homogeneous ground cover.

As shown in Figure 7.5, the MLE solution already yields a very acceptable estimate after the first iteration. This estimate hardly evolves further, which justifies the use of a single iteration of the MLE algorithm to estimate $\mathbf{r}_\alpha$.

On the contrary, the EM algorithm requires a certain number of iterations before reaching an acceptable estimate. The overestimation of the initial value of the covariance matrix $\mathbf{R}_\alpha^{(0)}$ results in an overestimation of the covariance of $\boldsymbol{\alpha}$ (the two last terms of (7.48)) inducing an overestimation of $\mathbf{r}_\alpha^{(k+1)}$. The higher the initial value of $\mathbf{R}_\alpha$, the higher this overestimation. In that case, the algorithm requires more steps before reaching convergence.

7.7.4 Comparison of the scattering coefficient estimation

This section compares the performances of the different estimation methods for different scenarios. The geometric setup is the same as in the previous section.

7.7.4.1 Homogeneous ground cover

First, let us considered a homogeneous terrain, with scattering coefficients of 0dB. The resulting scattering coefficient maps are illustrated in Figure 7.6. The scattering coefficients obtained using

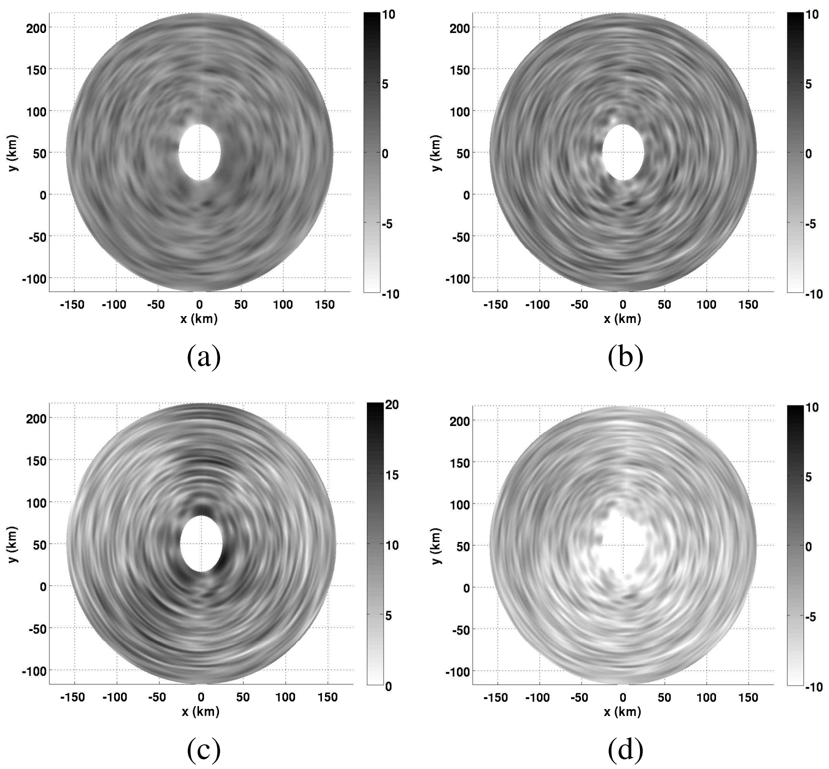


Figure 7.6: Estimated clutter map (homogeneous ground truth) (a) EM algorithm, (b) MLE, (c) CMF and (d) MAP estimator.

the MLE and the EM algorithm are relatively uniform. Both provide a very similar scattering coefficient map. The estimate obtained using the CMF exhibits spatial artifacts, which indicates that a constant correction factor does not completely compensate the bias. This result is also

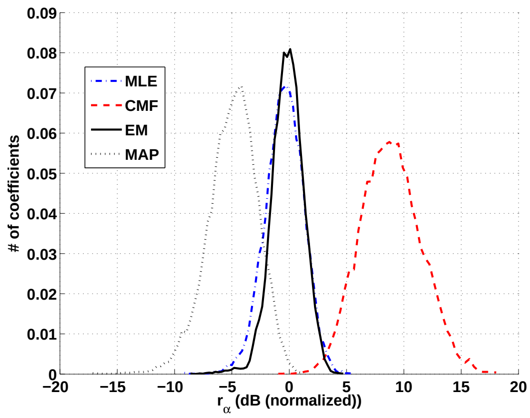


Figure 7.7: Histogram of the obtained scattering coefficients (homogeneous ground truth = 0dB).

visible in Figure 7.7 showing the histogram of the scattering coefficients obtained using the different methods. The histogram of the CMF is biased and also wider than the histogram of the MLE and the EM respectively. Finally, the MAP underestimates the scattering coefficients, as shown in Figure 7.6 and 7.7. Indeed, the covariance of $\boldsymbol{\alpha}$ is neglected in this method, which induces an underestimation of $\mathbf{r}_\alpha$.

7.7.4.2 Directional Tx antenna

This scenario, differing from the previous one, considers a sinc-shaped transmit antenna radiation diagram (without backlobe). The estimated scattering coefficients are shown in Figure 7.8. The scattering coefficients in the backlobe of the Tx antenna have no influence on the measured data and, thus, cannot be estimated. The values for those coefficients thus result from the regularization. A similar effect occurs for the coefficients located in the zeros of the antenna diagram, which is clearly visible in Figure 7.8 (a) where the values of the coefficients are partly imposed by the initial value of the covariance matrix. In the case of the MLE, the value of these coefficients is forced to zero by the use of the MAP estimate. The effect of the antenna diagram is also visible in Figure 7.8 (d) where the regularization forces to small values the coefficients which do not contribute to the received signal.

Furthermore, clearly, the CMF fails to provide a useful estimate. The CMF is directly affected by the antenna diagram through the factor c_i representing the geometric effects. By following a similar reasoning as in the previous sections, a similar expression to the CMF has been used [94] to obtain an estimate of $E\{|\alpha_i c_i|^2\}$.

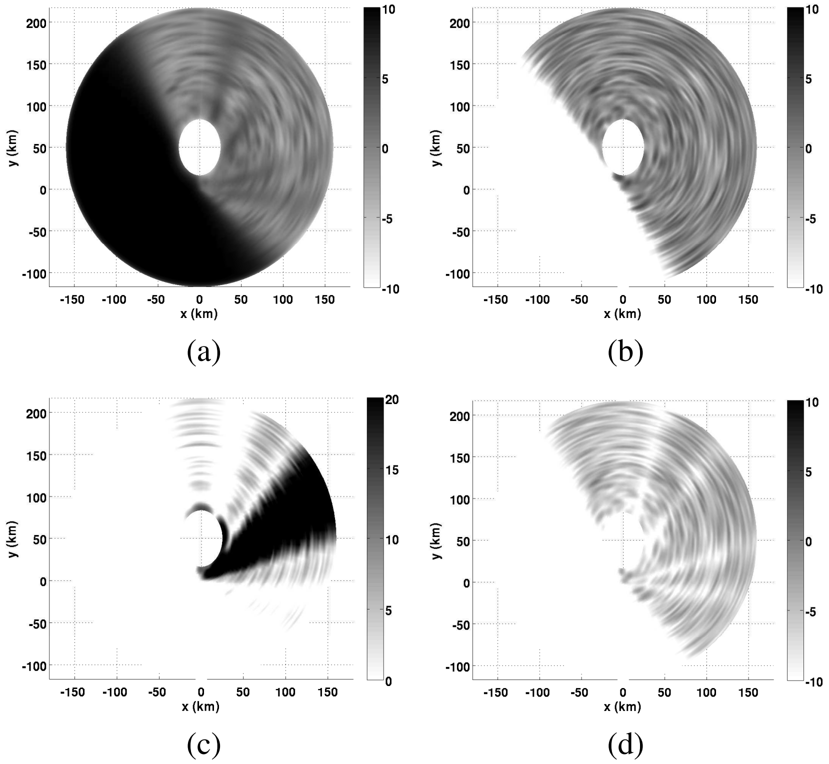


Figure 7.8: Estimated clutter map (directional Tx antenna diagram) (a) EM algorithm, (b) MLE, (c) CMF and (d) MAP estimator.

7.7.4.3 Non homogeneous ground cover

This new scenario considers non-uniform scattering coefficients with a difference of 10dB between the high and low scattering coefficients. This value is typically observed between (monostatic) ground backscattering and monostatic backscattering by the sea-surface in C-band [138], a demanding scenario in STAP. As illustrated in Figure 7.9, despite the non-uniformity, the pattern is still easily recognizable. Moreover, the histogram of Figure 7.10 is clearly bimodal with peaks around -5dB and +5dB, i.e., the exact values in the case of the EM and the MLE estimators. The MAP estimator underestimates the scattering coefficients. Although the pattern is still present in the CMF estimate, the corresponding histogram does not exhibit the bimodality, which denotes a large estimation error.

7.8 End to end performance

The end to end performance are assessed by comparing the SINR loss of the filters

$$\hat{\mathbf{w}} = k\hat{\mathbf{R}}^{-1}\mathbf{s} \tag{7.69}$$

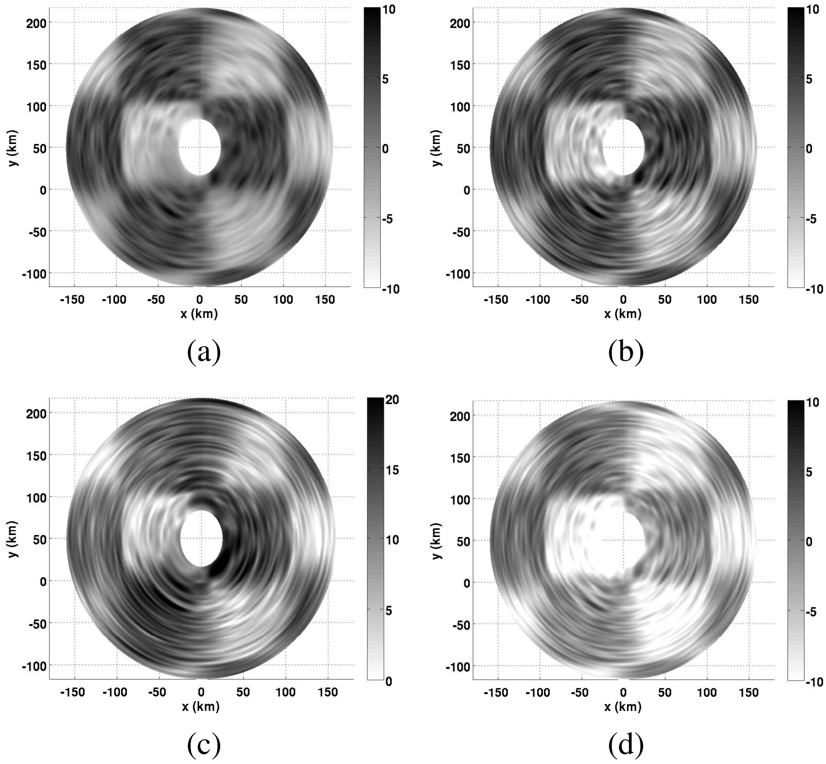


Figure 7.9: Estimated clutter map (checkerboard ground truth, i.e., non homogeneous) (a) EM algorithm, (b) MLE, (c) CMF and (d) MAP estimator.

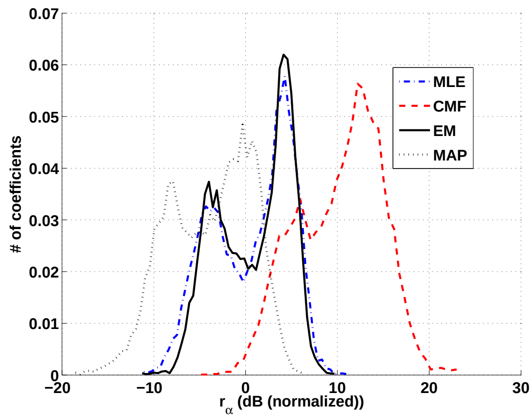
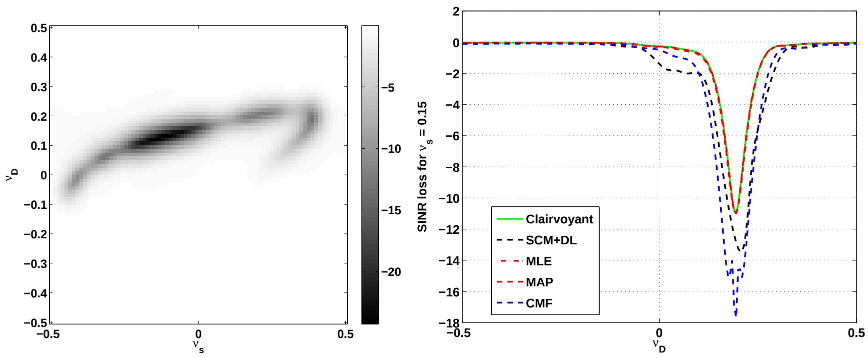


Figure 7.10: Histogram of the scattering coefficient (checkerboard patterned scattering coefficients with ±5dB).

where $\hat{\mathbf{R}}$ is the estimated covariance matrix obtained by replacing $\mathbf{R}_\alpha$ by the estimated scattering coefficients $\hat{\mathbf{R}}_\alpha$ in (7.5). The optimum filter, $\mathbf{w} = k\mathbf{R}^{-1}\mathbf{s}$ where $\mathbf{R}$ is the clairvoyant covariance matrix is used as reference. The scenario of interest is described in Section 7.7.2 and considers a homogeneous ground cover as well as a directional Tx antenna diagram. Figure 7.11 (a) shows



(a) (b)

Figure 7.11: SINR loss (a) due to the optimum filter and (b) comparison of SINR for the filter $\hat{\mathbf{w}}$ computed using the different estimates of $\mathbf{R}_\alpha$ (cut at $\nu_s = 0.15$).

the SINR loss of the optimum filter. As expected, the SINR loss is concentrated around the clutter power spectrum locus. The "modulation" of the SINR loss is due to the Tx antenna pattern, and in particular, as the backlobe signal is totally absent, it does not cause any loss. Figure 7.11 (b) shows a cut in the SINR loss for the filter $\hat{\mathbf{w}}$ computed using different estimates of $\mathbf{R}_\alpha$. The cut is taken at $\nu_s = 0.15$. The MLE and MAP methods prove to perform nearly as well as the optimum filter. Figure 7.12 presents the SINR losses with respect to the optimum filter, i.e., the *additional* SINR loss due to the use of an estimated clutter covariance matrix in place of the clairvoyant covariance matrix (see the first factor of (4.104)). The SINR loss, if the MLE method is used, is smaller than 0.5dB as this method actually provides an accurate estimate of the scattering coefficients. The underestimation of the scattering coefficients by the MAP method causes a larger SINR loss which however remains below 3dB. On the other hand, the CMF method, which overestimates the scattering coefficients causes a very large SINR loss, up to 15dB around the direction of the main beam of the Tx antenna, which clearly illustrates the need for an accurate estimation of the scattering coefficients. Finally, the SINR loss of the (diagonally loaded) sample matrix inversion is illustrated in Figure 7.12 (d) as a comparison.

7.9 Feasibility of joint estimation of the clutter map from several radars

If the scattering coefficients $\boldsymbol{\alpha}$ are isotropic, i.e. identical regardless of the look angle and the incidence angle, the formulation (7.3) can be extended to take into account observations of the same clutter patch from several radars, possibly at different ranges.

However, the isotropic assumption is far from being verified in practice. For instance, [12] reports up to 10dB azimuthal variations with sensors with a very coarse spatial resolution. A possible solution consists in taking into account the incidence angle effects by using for instance

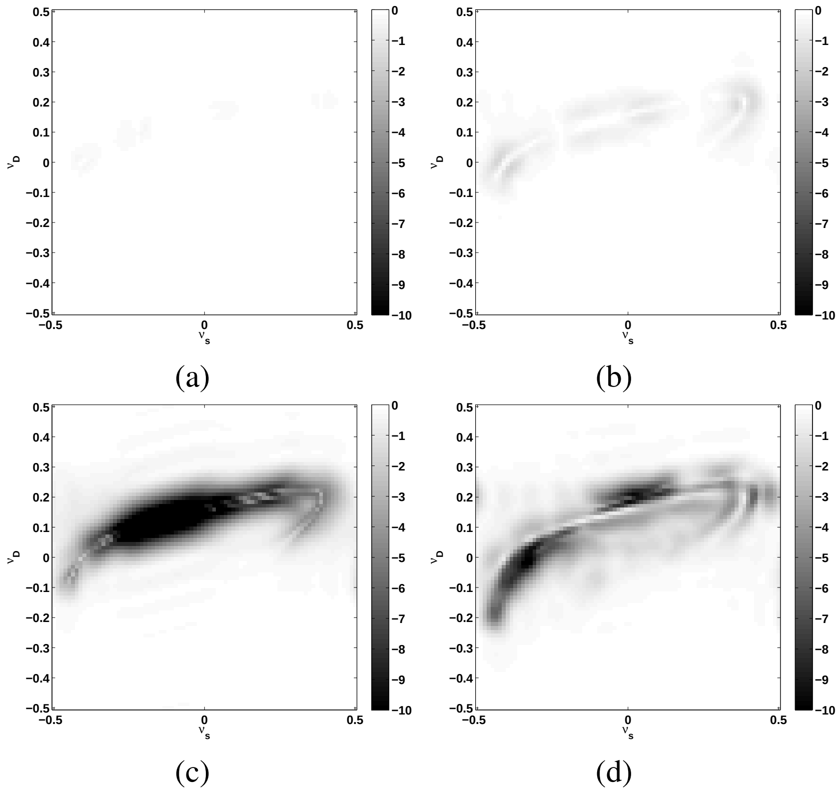


Figure 7.12: SINR losses with respect to the optimal filter if the scattering coefficients are estimated using (a) the MLE method, (b) the MAP method, (c) the crude matched filter method and (d) the loaded sample covariance matrix.

the model of Section 3.5.2.2 and by estimating the parameter γ_0 instead of the scattering coefficient σ_0. More generally speaking, one possibility to obtain isotropic quantities consists in using the underlying geophysical quantities. Over land, this could be the soil type, cover and moisture [131]. Over oceans, this could be the salinity, the wind speed and the wind direction [68, 69]. And over ice, this could be the ice type and the ice age [27, 35].

There are actual systems which use a sort of multistatic measurement[3] to derive some of these parameters. For instance the wind speed and the wind direction over oceans are operationally derived from measurements in three different directions [40, 69, 125]. Similarly, the ice information is routinely extracted from measurements in three different directions [28, 49]. However, given the complexity of these inversion algorithms [35, 69], the estimation of the geophysical parameters does not seem realistic in the context of the estimation of the scattering coefficients for moving target detection. Moreover, these models are typically highly non linear and the developments in this chapter assumed a linear model. The accuracy of the models used to perform the inversion is typically much lower than 0.1dB. As a pragmatic approach, approximated models might be considered since an accuracy of 1dB is probably reasonable in a context of moving target detection.

[3] Actually, several monostatic measurements with different look angles.

Chapter 8

Conclusions and perspectives

8.1 Summary and Conclusions

The present thesis has analyzed the issues in performing space-time adaptive processing with multistatic radars.

8.1.1 Signal modeling

The considered multistatic radar is composed of bistatic radars. The clutter is assumed Gaussian and hence entirely characterized by its covariance matrix. The clutter contributing to two different bistatic radars is inherently independent. Hence, the clutter covariance matrix of the multistatic radar can actually be decomposed in a combination of the clutter covariance matrices of each bistatic radar. A joint estimation of the different clutter covariance matrices is in principle feasible. Such a joint estimation would imply an estimation of the geophysical parameters based on existing models. These models are however either only approximative or very complex and require a demanding inversion process. Accordingly, taking into account the real-time objective of STAP, a multistatic clutter map estimation is unrealistic.

8.1.2 Multistatic target detection

This chapter presented the detection theory from a Neyman-Pearson point of view. It derived the test statistic for the Marcum and for the Swerling-I target models for bistatic radars. In both cases, the sufficient statistic is

$$|\mathbf{w}_p^\dagger \mathbf{y}_p|^2 \tag{8.1}$$

with $\mathbf{w}_p = k\mathbf{R}_p^{-1}\mathbf{s}_p$ which is the signal-dependent part of the test statistic. Extending this to multistatic radars shows that the test statistic in the case of the Swerling-I target model is the incoherent combination of the test statistic of each individual radar.

The performance of the detectors, evaluated first in a bistatic case, is extended further to the multistatic case. In a simplified multistatic case with Swerling-I target model, detection performance proved to increase if the test statistic of individual radars are combined. This effect is explained as the result of the diversity gain, as each individual radar observes a different realization of the random complex amplitude of the target echo, by increasing the chance for some of the radars to achieve a higher signal to noise ratio which ultimately benefits detection. For the more general case, the deflection coefficient is used as a measure of the “distance” between the distributions of the test statistic under both hypotheses, which allows to show that a multistatic radar is actually able to detect targets moving in any direction, which is not possible with individual monostatic or bistatic radars.

Next, the analysis of the impact of a GLRT detector, which uses the maximum likelihood estimate of the covariance matrix, makes possible the provision of a rigorous basis for the SINR loss metric commonly used to assess the loss in detection performance.

Finally, other, non optimum approaches to multistatic target detection are briefly discussed.

8.1.3 Clutter signal covariance matrix

This chapter made an analysis of the clutter covariance matrix in the spectral domain. In the case of uniform linear arrays (ULA), it is well known that the clutter covariance matrix power spectrum exhibits a so-called clutter ridge, i.e., the clutter energy is concentrated along a particular curve in the 2D spatio-temporal frequency domain. This curve, representing the so-called (2D) clutter power spectrum locus, is obtained by considering the limiting case of an infinite resolution (both in space and in time) continuous antenna. This curve is also known for having a very complex behavior as a function of the considered particular bistatic configuration. The concept of clutter power spectrum locus has been generalized to 4 dimensions, which corresponds to the general case of arbitrary antenna geometries. The effect of the bistatic configuration on the 4D clutter power spectrum locus are easier to interpret than in 2D. Moreover, it was shown that the 2D clutter power spectrum locus is actually a projection of the 4D curve, which explains the complex behavior of the 2D curve.

Further, this chapter examined the influence of the sampling, i.e., the spatial sampling due to the discrete nature of the receive elements and the temporal sampling due to the pulses transmitted by the transmitter (Chapter 6 generalize the temporal sampling). There is a direct relationship between the power spectrum estimated from the covariance matrix of a — discrete — signal and the power spectrum of the underlying continuous signal. This result has been used to establish the formal link between the clutter power spectrum locus and the power spectrum of the clutter covariance matrix.

The analysis of the range-dependence issue of the clutter statistics was made by analyzing the range-dependence of the clutter covariance matrix in the spectral domain. And this analysis, due to the direct link between the power spectrum of the covariance matrix and the clutter power spectrum locus, was made by examining the behavior of the clutter power spectrum locus as a function of the range. As a result, by a simple reasoning, it was possible to deduce the conditions under which there is no range dependence. In all generality, a range-independent power spectrum requires a static transmitter and a (possibly moving) receiver on the ground. This configuration has a useful application described in Chapter 6.

8.1.4 Transmitters of opportunity as signal source

This chapter has developed an optimum detector for signals of opportunity, which, in the absence of clutter signals, leads to the well-known ambiguity function, which is a characteristic of the waveform. The notion of ambiguity function has been generalized to the case where clutter is present at the range of interest. In which case, if the transmitted signal has a unit amplitude, the optimum filter is obtained by first demodulating the received signal and subsequently applying the usual generalized matched filter. A method proposed to speed-up the implementation consists in filtering and down-sampling the demodulated signal prior to performing the spectral analysis. This method has been extended to the multichannel case. As illustration, end-to-end results are shown and the feasibility of detection is demonstrated in the case of actual measurements with a GSM base-station as source of opportunity.

Target detection with signals of opportunity is limited due to the sidelobes of the considered ambiguity function. A CLEAN-like method to cope with this problem has been briefly discussed.

Finally, this chapter addresses issues arising from the extension to multistatic configurations.

8.1.5 Clutter map estimation

This chapter examines how to obtain a maximum likelihood estimate of the clutter plus noise covariance matrix needed to compute the filter that yields the detection statistic.

The main issues are the unavailability of enough identically distributed measurements in order that the sample covariance matrix provides an accurate estimation of the interference-plus-noise covariance matrix. Next, this chapter reviews the methods proposed in the literature to cope with the inhomogeneity of the scattering coefficients and the range dependence of the clutter power spectrum locus. Some of these methods can be interpreted as an estimation of the scattering coefficients. The most promising methods assume that the geometric configuration (i.e., position and velocity of the transmit and receive platforms) is known. In this case, the covariance matrix has a particular known structure and the problem can be recast as a structured covariance matrix estimation. A method is proposed to exploit this knowledge in a maximum likelihood framework in order to estimate the scattering coefficients.

Existing methods are shown to be approximations of the exact solution. A sensible and feasible approximation is proposed and provided accurate scattering coefficients estimates. End-to-end results are also given in the form of SINR losses. These results show again that the proposed approximation indeed provides excellent performance.

Finally, this chapter discusses the possibility to extend the method if multistatic measurements are available, and concludes that it is difficult to justify such extensions on theoretical bases.

8.2 Perspectives

The target detection theory developed in Chapter 4 assumes a Gaussian clutter model. Other models such as the spherical invariant random processes [165] also known as elliptical contour processes [60] should be considered to model the real-world clutter.

The covariance matrix estimation method presented in Chapter 7 assumes a known covariance model, which implies the perfect knowledge of the steering vectors. This might not always be the case. A first possible approach consists in using measured steering vectors. Alternatively, an on-line estimation of the parameters, possibly as an extension to [75], can also be attempted. The regularization implemented as a spatial averaging can be generalized. Indeed, whenever the clutter is homogeneous, spatially invariant spatial averaging is reasonable. However, in non-homogeneous clutter, spatially variant averaging taking into account the local contrast, such as is commonly done in SAR image processing [97] could be used. This can be seen as a generalization of the non-homogeneity detector where the non-homogeneous clutter range-ring is discarded while a spatially variant spatial filter would make it possible to discard only the non-homogeneous clutter patch.

The feasibility of STAP with transmitters of opportunity is demonstrated in the case of a bistatic radar, which could be extended to multiple transmitters of opportunity. This approach is perfectly feasible by using GSM base stations as, often, more than one GSM station is visible at any location. Alternatively, moving sources and a fixed receiver could be considered. Indeed, usual pulsed radar (Envisat, ERS, ...) could be used as transmitters of opportunity, however with a low revisit time and a poor ambiguity function. Signals from GPS and Galileo spacecrafts, having a noise-like characteristic, could also be envisaged. The low power of these signals is however a major challenge.

The measurements used in Chapter 6 were made with a static receiver. In order to fully demonstrate the feasibility of STAP with transmitters of opportunity, a moving receiver should also be considered. Given the discussion in Chapter 5, as long as the receiver is located on the ground, no range dependence of the covariance matrix is expected, which eases its estimation.

Coupling effects are neglected in this work. This assumption is probably overoptimistic and should be taken into account. A possible approach [1, 2] consists in modeling the coupling with the method of moments and in inverting the equation to obtain the incident field, in order compensating for this coupling.

Appendix A

Derivation of the isorange equation for bistatic configurations

In this section, the mathematical expression for the isorange intersection with the flat ground is derived.

The surfaces of equal round-trip delay are ellipsoids having Tx and Rx as focal points. These ellipsoids are centered in $(x_R/2, y_R/2, z_R/2)$ and have as semi-major axis $a = R/2$ and as semi-minor axis

$$b = \sqrt{\left(\frac{R}{2}\right)^2 - \left(\frac{R_{RT}}{2}\right)^2}$$

where R is the range under consideration.

Considering a coordinate system centered in the center of the ellipsoid and having an axis r' aligned with $\vec{RT}$, an y' axis perpendicular to the z' axis and parallel to the xy-plane and a z' axis perpendicular to the $r'y'$-plane, the ellipsoid thus has the following equation

$$\left(\frac{z'}{b}\right)^2 + \left(\frac{r'}{a}\right)^2 + \left(\frac{y'}{b}\right)^2 = 1.$$

The transformation of this coordinate system to a coordinate system centered in the center of the ellipsoid and having an axis r aligned with the projection of $\vec{RT}$ in the xy-plane, a y' axis perpendicular to the r-axis and lying in the xy-plane and a z axis perpendicular to the ry' plane are

$$\begin{cases} z' &= z\cos\theta - r\sin\theta \\ r' &= z\sin\theta + r\cos\theta \end{cases} \tag{A.1}$$

This yields the following equations for the ellipsoid

$$z^2\frac{a^2\cos^2\theta + b^2\sin^2\theta}{a^2b^2} + r^2\frac{a^2\sin^2\theta + b^2\cos^2\theta}{a^2b^2} + \frac{y'^2}{b^2} + rz\frac{(b^2-a^2)\sin\theta\cos\theta}{a^2b^2} = 1 \tag{A.2}$$

and the intersection with the plane $z = -H - z_R/2$ is

$$(H+z_R/2)^2\frac{a^2\cos^2\theta + b^2\sin^2\theta}{a^2b^2} + r^2\frac{a^2\sin^2\theta + b^2\cos^2\theta}{a^2b^2} + \frac{y'^2}{b^2} + r(H+z_R/2)\frac{(a^2-b^2)\sin\theta\cos\theta}{a^2b^2} = 1 \tag{A.3}$$

or adding and subtracting a term

$$(H + z_R/2)\frac{(a^2 - b^2)^2 \sin^2\theta \cos^2\theta}{a^2 b^2 (a^2 \sin\theta + b^2 \cos\theta)}$$

to obtain a perfect square expression yields

$$\frac{(r + r_0)^2}{A_1^2 F} + \frac{y'^2}{b^2 F} = 1 \tag{A.4}$$

where

$$r_0 = \frac{(H + z_R/2)(a^2 - b^2)\cos\theta \sin\theta}{a^2 \sin^2\theta + b^2 \cos^2\theta}, \tag{A.5}$$

$$F = 1 - \frac{(H + z_R/2)^2((a^2\cos^2\theta + b^2\sin^2\theta)(a^2\sin^2\theta + b^2\sin^2\theta) - (a^2 - b^2)\cos^2\theta\sin^2\theta)}{a^2 b^2 (a^2\sin^2\theta + b^2\cos^2\theta)} \tag{A.6}$$

and

$$A_1^2 = \frac{a^2 b^2}{a^2 \sin^2\theta + b^2\cos^2\theta} \tag{A.7}$$

the expression of this ellipse in the Cartesian coordinates xy is readily obtained using the transformation

$$\begin{cases} x & = r\cos\phi - y'\sin\phi \\ y & = r\sin\phi + y'\cos\phi \end{cases} \tag{A.8}$$

If the ground is assumed flat, (A.4) thus provides the expression of the isorange on the ground.

Appendix B

Useful distributions

B.1 Central chi-squared distribution

A central chi-squared distribution with ν degrees of freedom, denoted χ^2_ν, is defined as $x = \sum_{i=1}^{\nu} x_i^2$ where $x_i \sim \mathcal{N}(0,1)$ and the x_i are independent of each other [82, Chapter 2]. Its PDF is [82, Chapter 2]

$$p(x) = \begin{cases} \dfrac{1}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})} x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}} & x > 0 \\ 0 & x < 0 \end{cases}. \tag{B.1}$$

Its mean and variance are

$$E\{x\} = \nu \tag{B.2}$$
$$\mathrm{var}\{x\} = 2\nu. \tag{B.3}$$

The complement cumulative density function $Q_{\chi^2_\nu}(\eta)$ for even ν is given by

$$Q_{\chi^2_\nu}(\eta) = 1 - F_{\chi^2_\nu}(\eta) = \int_\eta^\infty p(x)dx = e^{-\frac{x}{2}} \sum_{k=0}^{\frac{\nu}{2}-1} \frac{1}{k!}(\frac{x}{2})^k \qquad \nu \geq 2 \tag{B.4}$$

where $F_{\chi^2_\nu}(\eta)$ is the cumulative density function.

An interesting, particular case arises when $\nu = 2$. In this case, the distribution is known as a negative exponential distribution with PDF

$$p(x) = \begin{cases} \dfrac{1}{2} e^{-\frac{x}{2}} & x > 0 \\ 0 & x < 0 \end{cases} \tag{B.5}$$

B.2 Non-central chi-squared distribution

A non-central chi-squared distribution with ν degrees of freedom and non-centrality parameter λ, denoted $\chi'^2_\nu(\lambda)$, is defined as $x = \sum_{i=1}^{\nu} x_i^2$ where $x_i \sim \mathcal{N}(\mu_i, 1)$ and the x_i are independent from each other [79, Chapter 28] and [82, Chapter 2]. The non centrality parameter λ is defined as $\lambda = \sum_{i=1}^{\nu} \mu_i^2$. Its PDF is defined by [82, Eq. (2.12)]

$$p(x) = \begin{cases} \frac{1}{2}\left(\frac{x}{\lambda}\right)^{\frac{\nu-2}{4}} e^{-\frac{x+\lambda}{2}} I_{\frac{\nu}{2}-1}(\sqrt{\lambda x}) & x > 0 \\ 0 & x < 0 \end{cases}. \tag{B.6}$$

where $I_r(u)$ is the modified Bessel function of the first kind and order r. Its mean and variance are

$$E\{x\} = \nu + \lambda \tag{B.7}$$

$$\mathrm{var}\{x\} = 2\nu + 4\lambda. \tag{B.8}$$

The non-central chi-squared cumulative density function (CDF) is given by [79, Chapter 28, eq. (5)]

$$F_{\chi_\nu'^2(\lambda)}(x) = \sum_{j=0}^{\infty} \frac{\lambda^j}{2^j j!} e^{-\frac{\lambda}{2}} F_{\chi^2_{\nu+2j}}(x). \tag{B.9}$$

Appendix C

Notable relations

C.1 Derivative of $\mathrm{Re}(\alpha a)$ and $|\alpha|^2$

As noted in [66, 81], the complex derivative of a real-valued function of a complex variable can be quite tricky and unintuitive, and is its useful to provide here particular results used in the text. Without repeating the discussion in [81], let us simply mention that the reason for this trickiness is the fact that a real-valued function of complex variables is not an analytic function.

Consider a real function $L(\alpha)$ of some complex parameter $\alpha = \alpha_R + j\alpha_I$. The definition of [66] is used here

$$\frac{\partial L}{\partial \alpha} = \frac{\partial L}{\partial \alpha_R} + j\frac{\partial L}{\partial \alpha_I} \tag{C.1}$$

and gives results that are often intuitive with respect to the corresponding real case. Note however that another definition is given in [81].

Let

$$L = \mathrm{Re}(\alpha a) \tag{C.2}$$

where α and a are complex numbers and let us compute $\frac{\partial L}{\partial \alpha}$. We thus have

$$L = \mathrm{Re}(\alpha a) = \alpha_R a_R - \alpha_I a_I \tag{C.3}$$

and hence

$$\frac{\partial L}{\partial \alpha} = \frac{\partial L}{\partial \alpha_R} + j\frac{\partial L}{\partial \alpha_I} = a_R - j a_I = a^*. \tag{C.4}$$

Similarly, for

$$L = |\alpha|^2 \tag{C.5}$$

one has

$$L = \alpha_R^2 + \alpha_I^2 \tag{C.6}$$

and hence

$$\frac{\partial L}{\partial \alpha} = 2\alpha \tag{C.7}$$

which is consistent with what would have been obtained for real a α.

C.2 Derivation of the maximum likelihood of a deterministic unknown α

In this section, the detailed derivation of the maximum likelihood estimation of α given the measurements

$$\mathbf{y} = \alpha\mathbf{s} + \mathbf{x}_{i+n} \tag{C.8}$$

is provided. Let us recall that $\mathbf{x}_{i+n}$ is assumed Gaussian with covariance matrix $\mathbf{R} = E\{\mathbf{x}_{i+n}\mathbf{x}_{i+n}^{\dagger}\}$.

The likelihood of H_1 assuming α deterministic but unknown is

$$L(H_1) = p(\mathbf{y}; \alpha, H_1) \tag{C.9}$$

where $p(\mathbf{y}; \alpha, H_1)$ is given by (4.4)

$$p(\mathbf{y}; \alpha, H_1) = \frac{1}{\pi^N \|\mathbf{R}\|} e^{-(\mathbf{y}-\alpha\mathbf{s})^{\dagger}\mathbf{R}^{-1}(\mathbf{y}-\alpha\mathbf{s})}. \tag{C.10}$$

The likelihood $L(H_1)$ and the log-likelihood $l(H_1) = \ln L(H_1)$ have the same maximum. The log-likelihood can thus be used, yielding

$$\begin{aligned} l(H_1) &= \ln \frac{1}{\pi^N \|\mathbf{R}\|} - (\mathbf{y} - \alpha\mathbf{s})^{\dagger}\mathbf{R}^{-1}(\mathbf{y} - \alpha\mathbf{s}) \\ &= \ln \frac{1}{\pi^N \|\mathbf{R}\|} - \mathbf{y}^{\dagger}\mathbf{R}^{-1}\mathbf{y} + 2\mathrm{Re}(\alpha\mathbf{y}^{\dagger}\mathbf{R}^{-1}\mathbf{s}) - |\alpha|^2\mathbf{s}^{\dagger}\mathbf{R}^{-1}\mathbf{s}. \end{aligned} \tag{C.11}$$

A necessary condition for α to be the maximum of this function is that its derivative be equal to zero

$$\frac{\partial l(H_1)}{\partial \alpha} = 2(\mathbf{y}^{\dagger}\mathbf{R}^{-1}\mathbf{s})^{\dagger} - 2\alpha\mathbf{s}^{\dagger}\mathbf{R}^{-1}\mathbf{s} = 0 \tag{C.12}$$

where the results of section C.1 was used. Hence the maximum likelihood estimate of α is given by

$$\hat{\alpha} = \frac{\mathbf{s}^{\dagger}\mathbf{R}^{-1}\mathbf{y}}{\mathbf{s}^{\dagger}\mathbf{R}^{-1}\mathbf{s}}. \tag{C.13}$$

One verifies easily that the second derivative is always negative and hence that $\hat{\alpha}$ is indeed the maximum.

C.3 Computation of the likelihood $L(H_1)$ when α is random

The likelihood of H_1 assuming α random with known PDF (4.13)

$$p(\alpha) = \frac{1}{\pi\sigma_{\alpha}^2} e^{-\frac{|\alpha|^2}{\sigma_{\alpha}^2}}. \tag{C.14}$$

is

$$L(H_1) = p(\mathbf{y}; H_1) = \int_D p(\mathbf{y}|\alpha; H_1)p(\alpha)d\alpha \tag{C.15}$$

where the domain of the integral D is $\mathbb{C}$ and $p(\mathbf{y}|\alpha; H_1)$ is given by (4.4)

$$p(\mathbf{y}|\alpha; H_1) = \frac{1}{\pi^N \|\mathbf{R}\|} e^{-(\mathbf{y}-\alpha\mathbf{s})^\dagger \mathbf{R}^{-1}(\mathbf{y}-\alpha\mathbf{s})}. \tag{C.16}$$

Hence

$$L(H_1) = \int_D p(\mathbf{y}|\alpha; H_1)p(\alpha)d\alpha \tag{C.17}$$

$$= \frac{1}{\pi^N \|\mathbf{R}\|} \frac{1}{\pi\sigma_\alpha^2} \int_D e^{-\mathbf{y}^\dagger \mathbf{R}^{-1}\mathbf{y} + \alpha\mathbf{y}^\dagger \mathbf{R}^{-1}\mathbf{s} + \alpha^* \mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{y} + |\alpha|^2 (\frac{1}{\sigma_\alpha^2} + \mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{s})} d\alpha \tag{C.18}$$

completing the square yields

$$L(H_1) = \frac{1}{\pi^N \|\mathbf{R}\|} \frac{1}{\pi\sigma_\alpha^2} e^{-\mathbf{y}^\dagger \mathbf{R}^{-1}\mathbf{y}} e^{\frac{|\mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{y}|^2}{\frac{1}{\sigma_\alpha^2} + \mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{s}}} \int_D e^{-\left|\alpha\sqrt{\frac{1}{\sigma_\alpha^2} + \mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{s}} - \frac{\mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{y}}{\sqrt{\frac{1}{\sigma_\alpha^2} + \mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{s}}}\right|^2} d\alpha \tag{C.19}$$

and since [38]

$$\int_D e^{-|\beta|^2} d\beta = \pi \tag{C.20}$$

the integral in (C.19) yields

$$\int_D e^{-\left|\alpha\sqrt{\frac{1}{\sigma_\alpha^2} + \mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{s}} - \frac{\mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{y}}{\sqrt{\frac{1}{\sigma_\alpha^2} + \mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{s}}}\right|^2} d\alpha = \frac{\pi}{\sqrt{\frac{1}{\sigma_\alpha^2} + \mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{s}}} \tag{C.21}$$

and finally

$$L(H_1) = \frac{1}{\pi^N \|\mathbf{R}\|} \frac{1}{\pi\sigma_\alpha^2} \frac{\pi}{\sqrt{\frac{1}{\sigma_\alpha^2} + \mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{s}}} e^{-\mathbf{y}^\dagger \mathbf{R}^{-1}\mathbf{y}} e^{\frac{|\mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{y}|^2}{\frac{1}{\sigma_\alpha^2} + \mathbf{s}^\dagger \mathbf{R}^{-1}\mathbf{s}}}. \tag{C.22}$$

Appendix D

Alternative expression of the MAP estimate of α

D.1 Matrix inversion lemma

The matrix inversion lemma states

$$(\mathbf{A} + \mathbf{B}\mathbf{C}\mathbf{D})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{D}\mathbf{A}^{-1}\mathbf{B} + \mathbf{C}^{-1})^{-1}\mathbf{D}\mathbf{A}^{-1} \tag{D.1}$$

where the matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ have appropriate dimensions and the required inverses are assumed to exist.

D.2 Alternative expression of the MAP estimate of α

There are two possible expressions for the MAP estimate of $\boldsymbol{\alpha}$. Showing that these expressions are equivalent is a classical exercise in estimation theory. However, it is useful to have these two expression since they can be used to gain insight in some of the estimators developed in Section 7.5.

The MAP estimate of $\boldsymbol{\alpha}$ is given by (7.22)

$$\hat{\boldsymbol{\alpha}} = (\mathbf{S}^\dagger \mathbf{R}_n^{-1} \mathbf{S} + \hat{\mathbf{R}}_\alpha^{-1})^{-1} \mathbf{S}^\dagger \mathbf{R}_n^{-1} \mathbf{y} \tag{D.2}$$

where $\hat{\mathbf{R}}_\alpha$ is the matrix providing the prior knowledge about $\boldsymbol{\alpha}$.

We successively have

$$\begin{aligned}
\hat{\boldsymbol{\alpha}} &= (\mathbf{S}^\dagger \mathbf{R}_n^{-1} \mathbf{S} + \hat{\mathbf{R}}_\alpha^{-1})^{-1} \mathbf{S}^\dagger \mathbf{R}_n^{-1} \mathbf{y} && \text{(D.3)}\\
&= \hat{\mathbf{R}}_\alpha (\hat{\mathbf{R}}_\alpha^{-1} + \mathbf{S}^\dagger \mathbf{R}_n^{-1} \mathbf{S} - \mathbf{S}^\dagger \mathbf{R}_n^{-1} \mathbf{S})(\mathbf{S}^\dagger \mathbf{R}_n^{-1} \mathbf{S} + \hat{\mathbf{R}}_\alpha^{-1})^{-1} \mathbf{S}^\dagger \mathbf{R}_n^{-1} \mathbf{y} && \text{(D.4)}\\
&= \hat{\mathbf{R}}_\alpha [\mathbf{I} - \mathbf{S}^\dagger \mathbf{R}_n^{-1} \mathbf{S}(\mathbf{S}^\dagger \mathbf{R}_n^{-1} \mathbf{S} + \hat{\mathbf{R}}_\alpha^{-1})^{-1}] \mathbf{S}^\dagger \mathbf{R}_n^{-1} \mathbf{y} && \text{(D.5)}\\
&= \hat{\mathbf{R}}_\alpha \mathbf{S}^\dagger [\mathbf{R}_n^{-1} - \mathbf{R}_n^{-1} \mathbf{S}(\mathbf{S}^\dagger \mathbf{R}_n^{-1} \mathbf{S} + \hat{\mathbf{R}}_\alpha^{-1})^{-1} \mathbf{S}^\dagger \mathbf{R}_n^{-1}] \mathbf{y} && \text{(D.6)}\\
&= \hat{\mathbf{R}}_\alpha \mathbf{S}^\dagger (\mathbf{S} \mathbf{R}_\alpha \mathbf{S}^\dagger + \mathbf{R}_n)^{-1} \mathbf{y} && \text{(D.7)}
\end{aligned}$$

where the matrix inversion lemma was used to go from (D.6) to (D.7), and (D.7) is equivalent to (7.46). In both cases, the current available *a priori* estimate of the covariance matrix is used. In the case of the MAP estimator (D.7), the *a priori* covariance matrix is used and this covariance matrix denotes the *a priori* knowledge while, in (7.46), the estimate obtained at the previous iteration is used.

Bibliography

[1] R. S. Adve and T. K. Sarkar. Elimination of the effects of mutual coupling in an adaptive nulling system with a look direction constraint. In *Proc. of the IEEE Antennas and Propagation Society International Symposium*, pages 1164–1167, May 1996.

[2] R. S. Adve, T. K. Sarkar, and R. A. Schneible. Accounting for the effects of mutual coupling in adaptive antennas. In *Proc. of the IEEE National Radar Conference*, Syracuse, NY, May 1997.

[3] R. S. Adve, R. A. Schneible, G. Genello, and P. Antonik. Waveform-space-time adaptive processing for distributed aperture radars. In *Proc. of the International IEEE Radar Conference 2005*, pages 93–97, May 2005.

[4] R. S. Adve, R. A. Schneible, and R. McMillan. Adaptive space/frequency processing for distributed aperture radars. In *Proc. of the International IEEE Radar Conference 2003*, pages 160–164, Sept. 2003.

[5] R. S. Adve and M. C. Wicks. Joint domain localized adaptive processing using measured steering vectors. In *Proc. of the IEEE Radar Conference 1998*, pages 165–170, Dallas, TX, May 1998.

[6] E. Anterrieu, B. Picard, and A. Tanner. On the regularization of inverse problems in imaging radiometry by aperture synthesis applications. In *Proceedings of the 8th Specialist Meeting on Microwave Radiometry & Remote Sensing*, pages 64–67, Rome, IT, Feb. 2004.

[7] S. R. J. Axelsson. On the theory of noise Doppler radar. In *Proceedings of IGARSS 2000*, pages 856–860, July 2000.

[8] S. R. J. Axelsson. Noise radar for range/Doppler processing and digital beamforming using low-bit ADC. *IEEE Transactions on Geoscience and Remote Sensing*, 41(12):2703–2720, Dec. 2003.

[9] S. R. J. Axelsson. Noise radar using random phase and frequency modulation. *IEEE Transactions on Geoscience and Remote Sensing*, 42(11):2370–2384, Nov. 2004.

[10] R. Baltes and G. van Keuk. Tracking multiple manoeuvering targets in a network of passive radars. In *Record of the International IEEE Radar Conference 1995*, pages 304–309, May 1995.

[11] M. Barkat. *Signal detection and estimation*. Artech house, Norwood, MA, 2005.

[12] Z. Bartalis, K. Scipal, and W. Wagner. Azimuthal anisotropy of scatterometer measurements over land. *IEEE Transactions on Geoscience and Remote Sensing*, 44(8):2083–2092, Aug. 2006. doi:10.1109/TGRS.2006.872084.

[13] T. A. Barton and S. T. Smith. Structured covariance estimation for space-time adaptive processing. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 3493–3496, Minneapolis, MN, Apr. 1997.

[14] D. Baumgarten. Optimum detection and receiver performance for multistatic radar configurations. In *Proceedings of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP'82)*, pages 359–362, 1982.

[15] W. Beide. The nature of bistatic and multistatic radar. In *Proceedings of the 2001 CIE Int. Conf. on Radar*, pages 882–884, 2001.

[16] D. C. Bell and R. M. Narayanan. Theoretical aspects of radar imaging using stochastic waveforms. *IEEE Transactions on Signal Processing*, 49(2):394–400, Feb. 2001.

[17] J. S. Bergin, P. M. Techau, J. E. Don Carlos, and J. R. Guerci. Radar waveform optimization for colored noise mitigation. In *Proc. of the International IEEE Radar Conference 2005*, pages 149–154, Arlington, VA, May 2005.

[18] S. D. Blunt and K. Gerlach. Aspects of multistatic adaptive pulse compression. In *Proc. of the International IEEE Radar Conference 2005*, pages 104–108, May 2005.

[19] G. K. Borsari. Mitigating effects on STAP processing caused by an inclined array. In *Proceedings of the IEEE Radar Conference 1998*, pages 135–140, Dallas, TX, May 1998.

[20] I. Bradaric, G. Capraro, D. D. Weiner, and M. C. Wicks. Multistatic radar systems signal processing. In *Proceedings of the IEEE Radar Conference*, pages 106–113, Verona, NY, Apr. 2006.

[21] L. E. Brennan and J. D. Mallett. Efficient simulation of external noise incident on arrays. *IEEE Transactions on Aerospace and Electronic Systems*, 25(6):740–741, Sept. 1976.

[22] L. E. Brennan and L. S. Reed. Theory of adaptive radar. *IEEE Transactions on Aerospace and Electronic Systems*, 9(2):237–252, Mar. 1973.

[23] Y. Bresler. Maximum likelihood estimation of a linearly structured covariance with application to antenna array processing. In *Proc. Fourth Annual ASSP Workshop on Spectral Estimation and Modeling*, pages 172–175, Aug. 1998.

[24] D. Bruyère and N. A. Goodman. Performance of multistatic space-time adaptive processing. In *Proceedings of the IEEE Radar Conference*, pages 533–538, Verona, NY, Apr. 2006.

[25] J. P. Burg, D. G. Luenberger, and D. L. Wenger. Estimation of structured covariance matrices. *Proceedings of the IEEE*, 70(9):963–974, Sept. 1982.

[26] J. Capon. High-resolution frequency-wavenumber spectrum analysis. *Proceedings of the IEEE*, 57(8):1408–1419, Aug. 1969.

[27] A. Cavanié. An empirical C-band backscatter model over arctic sea ice from ERS-1 AMI-wind data. In *Proc. of a Joint ESA-Eumetsat Workshop on Emerging Scatterometer Applications*, pages 99–106, Noordwijk, The Netherlands, Nov. 1995. ESTEC.

[28] A. Cavanié, F. Gohin, Y. Quilfen, and P. Lecomte. Identification of sea ice zones using the AMI-wind: Physical bases and applications to the FDP and CERSAT processing chains. In *Proc. Second ERS-1 Validation Workshop*, pages 1009–1012, Hamburg, Germany, Oct. 1993. ESTEC.

[29] L. Cazzani, C. Colesanti, D. Leva, G. Nesti, C. Prati, F. Rocca, and D. Tarchi. A ground based parasitic sar experiment. In *Proceedings of IGARSS '99*, volume 3, pages 1525–1527, June 1999.

[30] C.-F. Chang and M. R. Bell. Image residue approach for masked targets detection and estimation. *IEEE Transactions on Signal Processing*, 53(8):2975–2983, Aug. 2005.

[31] M. Cherniakov, T. Zeng, and E. Plakidis. GALILEO signal based bistatic system for avalanche prediction. In *Proceedings of IGARSS '03*, pages 784–786, July 2003.

[32] E. Conte, E. D'Addio, A. Farina, and M. Longo. Multistatic radar detection: synthesis and comparison of optimum and suboptimum receivers. *IEE Proceedings, Part F - Communications, Radar and Signal Processing*, 130(6):484–494, Oct. 1983.

[33] J. C. Curlander and R. N. McDonough. *Synthetic Aperture Radar, Systems and Signal Processing*. John Wiley and Sons, NY, 1991.

[34] P. J. Davis. *Circulant Matrices*. John Wiley and Sons, NY, 1979.

[35] S. de Haan and A. Stoffelen. Ice discrimination using ERS scatterometer. Technical report, KNMI, Sept. 2001.

[36] W. P. Delaney. Changing world — changing nature of conflicts – a critical role for military radar. *IEEE Aerospace and Electronic Systems Magazine*, pages 8–14, Nov. 1995.

[37] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via de EM algorithm. *Journal of the Royal Statistical Society*, B39:1–37, 1977.

[38] H. B. Dwight. *Tables of integrals and other mathematical data*. Macmillan, 1957.

[39] A. Farina and A. Russo. Radar detection of correlated targets in clutter. *IEEE Transactions on Aerospace and Electronic Systems*, 22(5):513–532, Sept. 1986.

[40] J. Figa-Saldana, J. J. W. Wilson, E. Attema, R. Gelsthorpe, M. R. Drinkwater, and A. Stoffelen. The advanced scatterometer(ASCAT) on the meteorological operational (MetOp)'s platforms: A follow on for European wind scatterometers. *Can. J. Remote Sensing*, 28(3):404–412, 2002.

[41] E. Fishler, A. Haimovich, R. Blum, D. Chizhik, L. Cimini, and R. Valenzuela. MIMO radar: and idea whose time has come. In *Proceedings of the IEEE Radar Conference*, pages 71–76, Philadelphia, PE, Apr. 2004.

[42] E. Fishler, A. Haimovich, R. Blum, L. Cimini, D. Chizhik, and R. Valenzuela. Performance of MIMO radar systems: advantages of angular diversity. In *Conference Record of the Thirty-Eight Asilomar Conference on Signals, Systems and Computers*, volume 1, pages 305–309, Nov. 2004.

[43] E. Fishler, A. Haimovich, R. S. Blum, L. Cimini, D. Chizhik, and R. A. Valenzuela. Spatial diversity in radars — models and detection performance. *IEEE Transactions on Signal Processing*, 54(3):823–838, Mar. 2006.

[44] T. L. Foreman. Reinterpreting the CLEAN algorithm as an optimum detector. In *Proceedings of the IEEE Radar Conference*, pages 769–775, Verona, NY, Apr. 2006.

[45] G. Franceschetti and R. Lanari. *Synthetic aperture radar processing*. CRC Press, 1999.

[46] D. R. Fuhrmann. Application of Toeplitz covariance estimation to adaptive beamforming and detection. *IEEE Transactions on Signal Processing*, 39(10):2194–2198, Oct. 1991.

[47] D. R. Fuhrmann. Structured covariance estimation: Theory, application and recent results. In *IEEE Sensor Array and Multichannel Signal Processing Workshop*, Waltham, MA, July 2006.

[48] K. Gerlach, A. K. Shackelford, and S. D. Blunt. A novel approach to shared-spectrum multistatic radar. In *Proceedings of the IEEE Radar Conference*, pages 437–443, Verona, NY, Apr. 2006.

[49] F. Gohin and C. Maroni. ERS scatterometer polar sea ice grids: User manual. Technical Report C2-MUT-W-03-IF, IFREMER, Mar. 1998.

[50] J. S. Goldstein and I. S. Reed. A new method of Wiener filtering and its application to interference mitigation for communications. In *MILCOM'97 Proceedings*, volume 3, pages 1087–1091, Nov. 1997.

[51] J. S. Goldstein, I. S. Reed, and L. L. Scharf. A multistage representation of the Wiener filter based on orthogonal projections. *IEEE Transactions on Information Theory*, 44(7):2943–2959, Nov. 1998.

[52] R. C. Gonzalez and R. E. Woods. *Digital image processing*. Prentice Hall, Englewood Cliffs, NJ, 2008.

[53] N. A. Goodman and D. Bruyere. Optimum and decentralized detection for multistatic airborne radar. *IEEE Transactions on Aerospace and Electronic Systems*, 43(2):806–813, Apr. 2007.

[54] N. R. Goodman. Statistical analysis based on a certain multivariate complex gaussian distribution (an introduction). *Annals of Mathematical Statistics*, 34:152–177, 1963.

[55] S. D. Grève, P. Ries, F. D. Lapierre, and J. G. Verly. Framework and taxonomy for radar space-time adaptive processing (STAP) methods. *IEEE Transactions on Aerospace and Electronic Systems*, 43(3):1084–1098, July 2007.

[56] H. D. Griffiths and C. J. Baker. Measurement and analysis of ambiguity functions of passive radar transmissions. In *Proc. of the International IEEE Radar Conference 2005*, pages 321–325, May 2005.

[57] H. D. Griffiths, C. J. Baker, J. Baubert, N. Kitchen, and M. Treagust. Bistatic radar using satellite-borne illuminators. In *Proc. of RADAR 2002*, pages 1–5, Oct. 2002.

[58] H. D. Griffiths and N. R. W. Long. Television-based bistatic radar. *IEE Proceedings*, 133(7):649–657, Dec. 1986.

[59] J. R. Guerci. Theory and application of covariance matrix tapers for robust adaptive beamforming. *IEEE Transactions on Signal Processing*, 47(4):977–987, Apr. 1999.

[60] J. R. Guerci. *Space-Time Adaptive Processing for Radar*. Artech House, Norwood, MA, 2003.

[61] J. R. Guerci and J. S. Bergin. Principal components, covariance matrix tapers, and the subspace leakage problem. *IEEE Transactions on Aerospace and Electronic Systems*, 38(1):152–162, Jan. 2002.

[62] J. R. Guerci, J. S. Goldstein, and I. S. Reed. Optimal and adaptive reduced-rank STAP. *IEEE Transactions on Aerospace and Electronic Systems*, 36(2):647–663, Apr. 2000.

[63] H. Guo, K. Woodbridge, and C. Baker. Evaluation of WiFi beacon transmissions for wireless based passive radar. In *Proc. of the IEEE Radar Conference 2008*, May 2008.

[64] A. M. Haimovich and Y. Bar-Ness. An eigenanalysis interference canceler. *IEEE Transactions on Signal Processing*, 39(1):76–84, Jan. 1991.

[65] E. Hanle. Survey of bistatic and multistatic radar. *IEE Proceedings, Part F - Communications, Radar and Signal Processing*, 133(7):587–595, Dec. 1986.

[66] S. Haykin. *Adaptive Filter Theory*. Prentice Hall, Englewood Cliffs, NJ, 1986.

[67] C. W. Helstrom. *Elements of signal detection and estimation*. Prentice Hall, Englewood Cliffs, NJ, 1995.

[68] H. Hersbach. CMOD5 - an improved geophysical model function for ERS C-Band scatterometry. Technical Report 395, ECMWF, Jan. 2003.

[69] H. Hersbach, A. Stoffelen, and S. D. Haan. The improved C-band geophysical model function CMOD5. In *Proceedings of the Envisat Symposium*, Salzburg, Austria, Sept. 2004.

[70] R. K. Hersey, W. L. Melvin, and J. H. McClellan. Clutter-limited detection performance of multi-channel conformal arrays. *Signal Processing*, 84:1481–1500, May 2004.

[71] B. Himed. Effects of bistatic clutter dispersion on STAP systems. In *Proceedings of the IEEE Radar Conference 2002*, pages 360–364, Long Beach, CA, Apr. 2002.

[72] B. Himed and J. H. Michels. Performance analysis of the multi-stage Wiener filter. In *Records of the IEEE International Radar Conference*, pages 729–734, May 2000.

[73] B. Himed, Y. Zhang, and A. Hajjari. STAP with angle-Doppler compensation for bistatic airborne radars. In *Proceedings of the IEEE Radar Conference 2002*, pages 311–317, Long Beach, CA, Apr. 2002.

[74] J. Horner, K. Kubik, B. Mojarrabi, I. D. Longstaff, E. Donskoi, and M. Cherniakov. Passive bistatic radar sensing with LEOS based transmitters. In *Proceedings of IGARSS '02*, pages 438–440, June 2002.

[75] A. G. Jaffer, B. Himed, and P. T. Ho. Estimation of range-dependent clutter covariance by configuration system parameter estimation. In *Proceedings of the IEEE Radar Conference*, pages 596–601, Arlington, VA, May 2005.

[76] J. K. Jao. Coherent multilateral radar processing for precise target geolocation. In *Proceedings of the IEEE Radar Conference*, pages 325–330, Verona, NY, Apr. 2006.

[77] T. Johnsen, K. E. Olsen, S. Johnsrud, and R. Skjerpeng. Simultaneous use of multiple pseudo random noise codes in multistatic CW radar. In *Proceedings of IEEE Radar Conference*, pages 266–270, Philadelphia, PE, Apr. 2004.

[78] D. H. Johnson and D. E. Dudgeon. *Array Signal Processing – Concepts and Techniques*. Prentice-Hall, Englewood Cliffs, NJ, 1992.

[79] N. L. Johnson and S. Kotz. *Distributions in statistics — Continuous univariate distributions - 2*. John Wiley and Sons, NY, 1970.

[80] S. M. Kay. *Modern Spectral Estimation, Theory & Application*. Prentice-Hall, Englewood Cliffs, NJ, 1988.

[81] S. M. Kay. *Fundamentals of statistical signal processing — Estimation theory*. Prentice-Hall, Englewood Cliffs, NJ, 1993.

[82] S. M. Kay. *Fundamentals of statistical signal processing — Detection theory*. Prentice-Hall, Englewood Cliffs, NJ, 1998.

[83] D. R. Kirk, J. S. Bergin, P. M. Techau, and J. E. D. Carlos. Multi-static coherent sparse aperture approach to precision target detection and engagement. In *Proceedings of the IEEE Radar Conference*, Alexandria, VA, May 2005.

[84] R. Klemm. *Principles of space-time adaptive processing*. The Institution of Electrical Engineers (IEE), UK, 2002.

[85] R. Klemm. *Applications of Space-Time Adaptive processing*. The Institution of Electrical Engineers (IEE), UK, 2004.

[86] S. M. Kogon and M. A. Zatman. Bistatic STAP for airborne radar systems. In *Adaptive Sensor Array Processing (ASAP) Workshop*, MIT Lincoln Laboratory, Lexington, MA, Mar. 2000.

[87] M. Kubica, V. Kubica, X. Neyt, J. Raout, S. Roques, and M. Acheroy. Optimum target detection using emitters of opportunity. In *Proceedings of the IEEE Radar Conference*, pages 417–424, Verona, NY, Apr. 2006.

[88] K. Kulpa. Continuous wave radars – monostatic, multistatic and network. In *Advances in Sensing with security applications*, Il Ciocco, Italy, July 2005. NATO Advanced Study Institute.

[89] K. S. Kulpa and Z. Czekala. Ground clutter suppression in noise radar. In *RADAR 2004*, Toulouse, France, Oct. 2004.

[90] A. D. Lanterman. Radar imaging with variations of an expectation-maximization algorithm. In *Proceedings of the IEEE Workshop on Detection, Estimation, Classification, and Imaging*, page 51, Santa Fe, NM, Feb. 1999.

[91] A. D. Lanterman. Statistical imaging in radio astronomy via an expectation-maximization algorithm for structured covariance estimation. In J. A. O'Sullivan, editor, *Statistical Methods in Imaging: In Medicine, Optics, and Communication, a festschrift in honor of Donald L. Snyder's 65th birthday*. Springer-Verlag, Jan. 2000.

[92] F. D. Lapierre. *Registration-based Range-dependence Compensation in Airborne Bistatic Radar STAP*. PhD thesis, Université de Liège, Liège, BE, may 2005.

[93] F. D. Lapierre, X. Neyt, and J. G. Verly. Evaluation of a registration-based range-dependence compensation method for a bistatic STAP radar using simulated, random snapshots. In *Proc. RADAR'04*, Toulouse, France, Oct. 2004.

[94] F. D. Lapierre, P. Ries, and J. G. Verly. Computationally-efficient range-dependence compensation methods for bistatic radar STAP. In *Proceedings of the IEEE Radar Conference*, pages 714–719, Arlington, VA, May 2005.

[95] F. D. Lapierre and J. G. Verly. Registration-based solutions to the range-dependence problem in STAP radars. In *Adaptive Sensor Array Processing (ASAP) Workshop*, MIT Lincoln Laboratory, Lexington, MA, Mar. 2003.

[96] P. Lecomte. ERS wind scatterometer commissioning and in-flight calibration. In *Proceedings of a Joint ESA-Eumetsat Workshop on Emerging Scatterometer Applications – From Research to Operations*, pages 261–270, The Netherlands, Nov. 1998. ESTEC.

[97] J. S. Lee. Refined filtering of image noise using local statistics. *Computer Graphics and Image Processing*, 15:380–389, 1981.

[98] G. R. Legters. Using a plane-wave signal model to suppress airborne GMTI radar clutter and calibrate the array. In *Proceedings of KASSPER 03*, pages 283–288, Las Vegas, NV, Apr. 2003.

[99] G. R. Legters. Schism: The final fortnight. Technical report, SAIC, Satellite Beach, FL, Feb. 2005.

[100] G. R. Legters and J. R. Guerci. Physics-based airborne GMTI radar signal processing. In *Proceedings of the IEEE Radar conference 2004*, pages 283–288, Apr. 2004.

[101] N. Levanon. CW alternatives to the coherent pulse train-signals and processors. *IEEE Transactions on Aerospace and Electronic Systems*, 29(1):240–254, Jan. 1993.

[102] H. Li, P. Stoica, and J. Li. Computationally efficient maximum likelihood estimation of structured covariance matrices. *IEEE Transactions on Signal Processing*, 47(5):1314–1323, May 1999.

[103] Y. Li and C. H. Cheong. Block Toeplitz with Toeplitz block covariance matrix for space-time adaptive processing. In *Proc. of the IEEE Radar Conference 2002*, pages 363–368, 2002.

[104] S. Mallat. *A wavelet tour of signal processing*. Academic Press, San Diego, CA, USA, 1998.

[105] D. G. Manolakis, V. K. Ingle, and S. M. Kogon. *Statistical and Adaptive Signal Processing*. Artech House, Norwood, MA, 2005.

[106] F. Marvasti. *Nonuniform sampling: Theory and practice*. Kluwer academic, Dordrecht, NL, 2001.

[107] W. L. Melvin. Issues in bistatic STAP. In *Proceedings of the 32nd Southeastern Symposium on System Theory*, Tallahassee, FL, Mar. 2000.

[108] W. L. Melvin, M. J. Callahan, and M. E. Davis. Comparison of bistatic clutter mitigation algorithms for varying geometries. In *Proceedings of the IEEE International Radar Conference*, Arlington, VA, May 2005.

[109] W. L. Melvin, M. J. Callahan, and M. C. Wicks. Adaptive clutter cancellation in bistatic radar. In *Conference Record of the Thirty-Fourth Asilomar Conference on Signals, Systems and Computers, 2000*, volume 2, pages 1125–1130, Nov. 2000.

[110] W. L. Melvin, M. J. Callahan, and M. C. Wicks. Bistatic STAP: application to airborne radar. In *Proceedings of IEEE Radar Conference 2002*, pages 1–7, Long Beach, CA, Apr. 2002.

[111] W. L. Melvin, B. Himed, and M. E. Davis. Doubly adaptive bistatic clutter filtering. In *Proceedings of the IEEE Radar Conference 2003*, pages 171–178, Huntsville, AL, May 2003.

[112] W. L. Melvin and G. A. Showman. An approach to knowledge-aided covariance estimation. *IEEE Transactions on Aerospace and Electronic Systems*, 42(3):1021–1042, July 2006.

[113] W. L. Melvin and M. C. Wicks. Improving practical space-time adaptive radar. In *Proc. of the IEEE National Radar Conference 1997*, pages 48–53, Syracuse, NY, May 1997.

[114] W. L. Melvin, M. C. Wicks, and R. D. Brown. Assessment of multichannel airborne radar measurements for analysis and design of space-time processing architectures and algorithms. In *Proceedings of the 1996 IEEE National Radar Conference*, pages 130–135, May 1996.

[115] G. V. Morris. *Airborne pulsed Doppler radar*. Artech House, Norwood, MA, 1988.

[116] P. Moulin. A wavelet regularization method for diffuse radar-target imaging and speckle-noise reduction. *Journal of Mathematical Imaging and Vision*, 3:123–134, 1993.

[117] P. Moulin, J. A. O'Sullivan, and D. L. Snyder. A method of sieves for multiresolution spectrum estimation and radar imaging. *IEEE Transactions on Information Theory*, 38(2):801–813, Mar. 1992.

[118] R. M. Narayanan. Random noise monopulse radar system for covert tracking of targets. Technical report, University of Nebraska, Lincoln, NE, July 2002.

[119] R. M. Narayanan and M. Dawood. Doppler estimation using a coherent ultrawide-band random noise radar. *IEEE Transactions on Aerospace and Electronic Systems*, 48(6):868–878, June 2000.

[120] F. E. Nathanson. *Radar design principles*. McGraw Hill, NY, 1969.

[121] X. Neyt, M. Acheroy, and J. G. Verly. Maximum likelihood range dependence compensation for STAP. In *Proceedings of the IEEE Conference on Acoustics, Speech and Signal Processing*, Honolulu, HI, Apr. 2007.

[122] X. Neyt, M. Acheroy, and J. G. Verly. Registration-based range-dependence compensation for coherent multistatic STAP-based radar. In *Proceedings of the IEEE Radar Conference*, Waltham, MA, Apr. 2007.

[123] X. Neyt, P. Druyts, M. Acheroy, and J. G. Verly. Structured covariance matrix estimation for the range-dependent problem in STAP. In *Proceedings of the 4th IASTED International Conference on Antennas, Radar, and Wave Propagation*, Montreal, Canada, May 2007.

[124] X. Neyt, F. D. Lapierre, and J. G. Verly. Principle and evaluation of a registration-based range-dependence compensation method for STAP in case of arbitrary antenna patterns and simulated snapshots. In *Proc. Adaptive Sensor Array Processing (ASAP) Workshop*, MIT Lincoln Laboratory, Lexington, MA, Mar. 2004.

[125] X. Neyt, P. Pettiaux, and M. Acheroy. Scatterometer ground processing review for gyroless operations. In *Proceedings of SPIE Remote Sensing of the Ocean, Sea Ice and Large Water Regions 2002*, volume 4880, Crete, Greece, Sept. 2002.

[126] X. Neyt, J. Raout, M. Kubica, V. Kubica, S. Roques, J. G. Verly, and M. Acheroy. Feasibility of STAP for passive GSM-based radar. In *Proceedings of the IEEE Radar Conference*, pages 546–551, Verona, NY, Apr. 2006.

[127] X. Neyt, P. Ries, J. G. Verly, and F. D. Lapierre. Registration-based range-dependence compensation method for conformal-array STAP. In *Proc. Adaptive Sensor Array Processing (ASAP) Workshop*, MIT Lincoln Laboratory, Lexington, MA, June 2005.

[128] X. Neyt, J. G. Verly, and M. Acheroy. Range-dependence issues in multistatic STAP-based radar. In *Proceedings of the Fourth IEEE Workshop on Sensor Array and Multi-channel Processing (SAM'06)*, Waltham, MA, July 2006.

[129] H. N. Nguyen, J. D. Hiemstra, and J. S. Goldstein. The reduced rank multistage Wiener filter for circular array STAP. In *Proceedings of the IEEE Radar Conference 2003*, pages 66–70, Huntsville, AL, May 2003.

[130] F. C. Nicolls and G. de Jager. Maximum likelihood estimation of Toeplitz-block-Toeplitz covariances in the presence of subspace interference. In *Proceedings of the 14th International Conference on Pattern Recognition*, volume 2, pages 1595–1597, Brisbane, Australia, Aug. 1998.

[131] Y. Oh, K. Sarabandi, and F. T. Ulaby. An empirical model and an inversion technique for radar scattering from bare soil surfaces. *IEEE Transactions on Geoscience and Remote Sensing*, 30(2):370–381, Mar. 1992.

[132] C. Oliver and S. Quegan. *Understanding Synthetic Aperture Radar Images*. SciTech Publishing, Raleigh, NC, 2004.

[133] P. E. Pace. *Detecting and Classifying Low Probability of Intercept Radar*. Artech House, Norwood, MA, 2004.

[134] D. A. Page, B. Himed, and M. E. Davis. Improving STAP performance in bistatic space-based radar systems using an efficient expectation-maximization technique. In *Proc. of the IEEE Radar Conference 2005*, pages 109–114, May 2005.

[135] I. Papoutsis, C. J. Baker, and H. D. Griffiths. Netted radar and the ambiguity function. In *Proc. of the International IEEE Radar Conference 2005*, pages 883–888, May 2005.

[136] Y. Pawitan. *In all likelihood: Statistical modelling and inference using likelihood*. Oxford University Press, 2001.

[137] F. Pearson and G. Borsari. Simulation and analysis of adaptive interference suppression for bistatic surveillance radars. In *Adaptive Sensor Array Processing (ASAP) Workshop*, MIT Lincoln Laboratory, Lexington, MA, Mar. 2001.

[138] P. Pettiaux, X. Neyt, and M. Acheroy. Validation of the ERS-2 scatterometer ground processor upgrade. In *Proceedings of SPIE Remote Sensing of the Ocean, Sea Ice and Large Water Regions 2002*, volume 4880, Crete, Greece, Sept. 2002.

[139] S. U. Pillai, H. S. Oh, D. C. Youla, and J. R. Guerci. Optimal transmit-receiver design in the presence of signal-dependent interference and channel noise. *IEEE Transactions on Information Theory*, 46(2):577–584, Mar. 2000.

[140] C. Prati, F. Rocca, D. Giancola, and A. M. Guarnieri. Passive geosynchronous SAR system reusing backscattered digital audio broadcasting signals. *IEEE Transactions on Geoscience and Remote Sensing*, 36(6):1973–1976, Nov. 1998.

[141] D. J. Rabideau and P. Parker. Ubiquitous MIMO multifunction digital array radar. In *Conference Record of the Thirty-Seventh Asilomar Conference on Signals, Systems and Computers*, volume 1, pages 1057–1064, Nov. 2003.

[142] I. S. Reed, J. D. Mallett, and L. E. Brennan. Rapid convergence rate in adaptive arrays. *IEEE Transactions on Aerospace and Electronic Systems*, 10(6):853–863, Nov. 1974.

[143] D. W. Rieken and D. R. Fuhrmann. Generalizing MUSIC and MVDR for distributed arrays. In *Conference Record of the Thirty-Sixth Asilomar Conference on Signals, Systems and Computers*, volume 1, pages 861–865, Nov. 2002.

[144] D. W. Rieken and D. R. Fuhrmann. Generalizing MUSIC and MVDR for multiple noncoherent arrays. *IEEE Transactions on Signal Processing*, 52(9):2396–2406, Sept. 2004.

[145] P. Ries, X. Neyt, F. D. Lapierre, and J. G. Verly. Fundamental principles associated with spatial and Doppler frequencies for arbitrary measurement configurations and antenna arrays in radar STAP. *IEEE Transactions on Aerospace and Electronic Systems*, 2008. Accepted.

[146] A. W. Rihaczek. *Principles of High-Resolution Radar*. McGraw-Hill, NY, 1969.

[147] M. A. Ringer and G. J. Frazer. Waveform analysis of transmissions of opportunity for passive radar. In *Proc. of the Fifth International Symposium on Signal Processing and Its Applications*, pages 511–514, Brisbane, Australia, Aug. 1999.

[148] F. C. Robey, S. Coutts, D. Weikle, J. C. McHarg, and K. Cuomo. MIMO radar theory and experimental results. In *Conference Record of the Thirty-Eight Asilomar Conference on Signals, Systems and Computers*, volume 1, pages 300–304, Nov. 2004.

[149] F. C. Robey and D. R. Fuhrmann. Structured covariance estimation via maximal representation of convex sets. In *Proceedings of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP'91)*, volume 5, pages 3249–3252, Apr. 1991.

[150] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg. A CFAR adaptive matched filter detector. *IEEE Transactions on Aerospace and Electronic Systems*, 28(1):208–216, Jan. 1992.

[151] A. Rosenfeld and A. C. Kak. *Digital Picture Processing*, volume 1. Academic Press, San Diego, CA, USA, 2 edition, 1982.

[152] R. Saini, M. Cherniakov, and V. Lenive. Direct path interference suppression in bistatic system: DTV based radar. In *Proc. of the International Radar Conference 2003*, pages 309–314, Sept. 2003.

[153] T. K. Sarkar, J. Koh, R. S. Adve, M. C. Wicks, S. Choi, and M. S. Palma. A pragmatic approach to adaptive antennas. *IEEE Antennas and Propagation Magazine*, 42(2):39–55, Apr. 2000.

[154] L. L. Scharf. *Statistical signal processing — detection, estimation, and time series analysis*. Addison-Wesley, MA, 1990.

[155] L. L. Scharf. *Statistical Signal Processing – Detection, Estimation, and Time Series Analysis*. Addison Wesley, MA, 1991.

[156] Y. V. Shkvarko. Estimation of wavefield power distribution in the remotely sensed environment: Bayesian maximum entropy approach. *IEEE Transactions on Signal Processing*, 50(9):2333–2346, Sept. 2002.

[157] Y. V. Shkvarko. Unifying regularization and bayesian estimation methods for enhanced imaging with remotely sensed data — Part I: Theory. *IEEE Transactions on Geoscience and Remote Sensing*, 42(5):923–931, May 2004.

[158] M. Skolnik. *Radar handbook*. McGraw-Hill, NY, 1990.

[159] D. L. Snyder, J. A. O'Sullivan, and M. I. Miller. The use of maximum likelihood estimation for forming images of diffuse radar targets from delay-Doppler data. *IEEE Transactions on Information Theory*, 35(3):536–548, May 1989.

[160] R. Srinivasan, P. Sharma, and V. Malik. Distributed detection of Swerling targets. *IEE Proceedings, Part F - Communications, Radar and Signal Processing*, 133(7):624–629, Dec. 1986.

[161] S. Stein. Algorithms for ambiguity function processing. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 29(3):588–599, June 1981.

[162] H. Sun, D. K. P. Tan, and Y. Lu. Design and implementation of an experimental GSM based passive radar. In *Proc. of the International Radar Conference 2003*, pages 418–422, Sept. 2003.

[163] D. K. P. Tan, H. Sun, Y. Lu, and W. Liu. Feasibility analysis of GSM signal for passive radar. In *Proc. of the International IEEE Radar Conference 2003*, pages 425–430, May 2003.

[164] D. K. P. Tan, H. Sun, Y. Lui, M. Lesturgie, and H. Chan. Passive radar using global system for mobile communication signal: theory, implementation and measurements. *IEE Proceedings – Radar, Sonar and Navigation*, 152(3):116–123, June 2005.

[165] J. Tang and Z. Zhu. Compound gaussian process and its derived processes [radar clutter]. In *Proceedings of the IEEE National Aerospace and Electronics Conference*, volume 2, pages 1074–1079, July 1997.

[166] A. N. Tikhonov and V. Y. Arsenin. *Solutions of ill-posed problems*. John Wiley & Sons, 1977.

[167] T. Tsao, M. Slamani, P. Varshney, D. Weiner, H. Schwarzlander, and S. Borek. Ambiguity function for a bistatic radar. *IEEE Transactions on Aerospace and Electronic Systems*, 33(3):1041–1051, July 1997.

[168] M. J. Turmon and M. I. Miller. Maximum-likelihood estimation of complex sinusoids and Toeplitz covariances. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 42(5):1074–1086, May 1994.

[169] H. L. Van Trees. *Detection, estimation and modulation theory — Part I.* John Wiley and Sons, NY, 1968.

[170] H. L. Van Trees. *Detection, estimation and modulation theory — Part III.* John Wiley and Sons, NY, 1971.

[171] H. L. Van Trees. *Detection, estimation and modulation theory — Part IV: Optimum Array Processing*. John Wiley and Sons, NY, 2002.

[172] J. G. Verly. Stap radar. Course given in the quality of Belgium Francqui-chair holder, May 2003.

[173] J. Ward. Space-time adaptive processing for airborne radar. Technical Report 1015, MIT Lincoln Laboratory, Lexington, MA, Dec. 1994.

[174] N. J. Willis. *Bistatic Radar*. Artech House, Norwood, MA, 1991.

[175] N. J. Willis and H. D. Griffiths. *Advances in Bistatic Radar*. SciTech Publishing, Raleigh, NC, 2007.

[176] M. A. Zatman. Circular array STAP. *IEEE Transactions on Aerospace and Electronic Systems*, 36(2):510–517, Apr. 2000.

[177] M. A. Zatman. Performance analysis of the derivative based updating method. In *Adaptive Sensor Array Processing (ASAP) Workshop*, MIT Lincoln Laboratory, Lexington, MA, Apr. 2001.

[178] Y. Zhang and B. Himed. Effects of geometry on clutter characteristics of bistatic radars. In *Proceedings of the IEEE Radar Conference 2003*, pages 417–424, Huntsville, AL, May 2003.

[179] H. Zheng, F. Li, J. Luo, and J. Lu. Bistatic radar experiment based on FM broadcast transmitter. In *Proc. RADAR'04*, Toulouse, France, Oct. 2004.