In this study we are interested by the station design in term of parameter accuracy, and more particularly of the azimuth. In short, in presence of only pure coherent acoustic waves, it is clear that the best design is to locate sensors at far as possible. For which reason could we take into account an upper bound on the aperture ?

Regarding the size of the station and the distance of the source, a possible phenomena, which could induce a limitation on the station aperture in presence of a coherent acoustic signal, is that the loss of coherence (LOC) which usually increases when the distance between sensors increases.

For conducting this study the 3 points following program has been considered:

- choosing an index to evaluate the accuracy of the station. A commonly used index is the Cramer-Rao bound (CRB). Typically that depends on the geometry of the station, the level of noise and the LOC features. A summary of the CRB could be the area/volume of the confidence ellipsoid.

- choosing features to characterize a geometry. We have retained (i) the isotropy and (ii) the uniform distribution of the inter-distances. The isotropy is easy to check and also it is simple to correct if necessary by adding 2 sensors (resp. 3) for 2D station (resp. 3D station).

  We can not answer to the good effect or not of the inter-distance uniformity because that depends on the LOC model. Therefore we have to validate a such model.

- validating a LOC model.

  The used LOC feature is the magnitude square coherence (MSC). This index has the interesting property to be between 0 and 1, and equal to 1 iff the two signals are spatially coherent.

  The approach to determine the LOC model is based on the observation analysis. It is conducted as it follows: we consider the station I37 which consists of 10 sensors with 45 uniformly disibuted inter-distances. We base the analysis on the presence of a quasi-permanent coherent acoustical signal, saying the microbarom, in a frequency bandwidth large enough between 0.05 to 3 Hz.

  For a given frequency, we select the portions of signals where the MSC is greater than 0.8 on the three nearest sensors and study the decay of the MSC along the inter-distance values.

Remark 1 (on the noise) *Typical station aperture is 2 km. Noise is mainly due to the wind.*

- *Therefore we assume that the noises are spatially uncorrelated regarding the sensor inter-distances.*

- *On the other hand, we assume that the noise levels are identical on all sensors. Although that is not realistic, it is worth to notice that the noise level is not directly related to the inter-distances between sensors. It follows that, for the station design understanding, we can consider there is no loss of generalities to assume that.*

Remark 2 (on the coherent source) *To be able to study the LOC, we need a permanent source. The microbarom plays this role in the following.*

Remark 3 (on the isotropy) *In the absence of LOC, the station is isotrope iff*

$$XX^T \quad \propto \quad I_d \tag{1}$$

where X is $d \times M$ matrix whose the m-th colomun if the coordinates of the m-th sensor (in any system of coordinates).

It follows that given an arbitrary array with M sensors in $\mathbb{R}^d$ it is always possible to complete with d locations in such a way the new array is isotrope, i.e. verifies (1), see annex.

Remark 4 (on the geometry) *In the absence of LOC, we will see below that the best performances are associated with inter-distances as large as possible. In this case the maximization of the minimal distance leads to locate the sensors on a circle. But in this case the distribution of the distances is not uniform. It follows that in presence of LOC, several sensors are concern with the same LOC. It seems (it is not a proof) that a rule could be to locate the sensors in such a way that the distances are more uniform (see figure 1).*
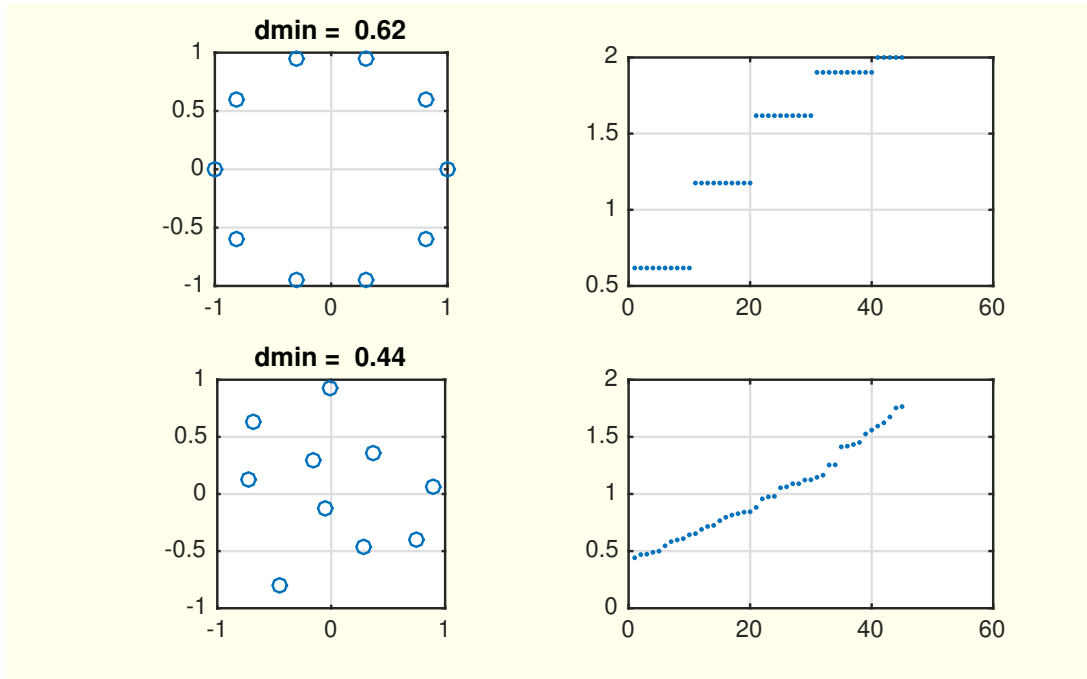
Figure 1: *Sensor locations with inter distances*

1 LOC model

1.1 Case of 2 sensors

By definition, two signals arriving on a two different locations are said non-coherent signals, or called noises, if they are spatially uncorrelated. More specifically if $s_1(t)$ and $s_2(t)$ denote the respective stationary signals observed in to different locations, the coherence level is defined by the magnitude square coherence:

$$\mathrm{MSC}(f) \;\; = \;\; \frac{|S_{12}(f)|^2}{S_{11}(f)S_{22}(f)}$$

where $S_{11}(f)$ and $S_{22}(f)$ denote the spectral densities of $s_1(t)$ and $s_2(t)$, and $S_{12}(f)$ the cross-spectrum. It is well known that $\mathrm{MSC}(f) \leq 1$.

Definition 1 (coherence) *When* $\mathrm{MSC}(f) = 1$, *we say that the 2 signals are perfectly coherent, elsewhere we speak about loss of coherence (LOC).*

It is shown that the $\mathrm{MSC}(f)$ between 2 signals is 1, if and only if it exists a filter, whose impulse respond is $g(t)$, such as $s_1(t) = g_1(t) \star s_2(t)$. A particular case is $g(t) = \delta(t - t_0)$ relative to a pure delay.

When $\mathrm{MSC}(f) = 0$ we say that the 2 signals are spatially uncorrelated and are called noises.

1.2 Case of M sensors

We consider a station with M sensors. There is only one acoustic source faraway from the station, in such a way it can be considered as planar wave. Therefore the M-ary signal writes:

$$x(t) \;\; = \;\; \underbrace{s(t;\theta)}_{\text{acoustic signal}} + \underbrace{w(t)}_{\text{noise}}$$

where θ denotes the 3D slowness vector. Under assumption of pure delay we can write for the m-th sensor located in r_m:

$$s_m(t;\theta) \;\; = \;\; s(t - r_m^T \theta) \tag{2}$$

It follows that, under the assumption that $s(t)$ is stationary random process with spectral density $\gamma_s(f)$, the spectral matrix of the M-ary process $s(t)$, whose the m-th entry is (2), writes:

$$\Gamma_s(f) \;\; = \;\; \gamma_s(f)d(f)d^H(f)$$

where $d(f)$ is a M-ary vector whose the m-entry writes $e^{-2j\pi f r_m^T \theta}$. Clearly the matrix $\Gamma_s(f)$ is of rank 1, and that is the general definition of the coherence, in accordance

4

with the definition 1 for only 2 sensors. In any case, the matrix $\Gamma_s(f) \geq 0$ for any frequency f.

If we assume that the spectral matrix of the noise $w(t)$ is $\sigma^2 I_M$ for any frequency, it follows that the MSC between any two sensors writes:

$$\text{MSC}_{\ell,\ell'}(f) = \frac{\gamma_s^2(f)}{(\gamma_s(f) + \sigma^2)(\gamma_s(f) + \sigma^2)}$$

Therefore the MSC level is not 1, because the noise.

Definition 2 (coherence) *M stationary signals are said spatially coherent if their spectral matrix is of rank 1, elsewhere we speak about LOC.*

Remark 5 *Coherence is not only restricted to the case of pure delays. Perfect coherence is also verified for filtering signals.*

1.3 Discrete domain

All signals are real and sampled at the sampling frequency $F_s = 20$ Hz. After sampling we obtain $x_n = x(n/F_s)$. For $k = 0$ to $(N - 1)$ we consider the M-ary discrete Fourier transform:

$$X_k = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x_n \, e^{-2j\pi n f_k} \quad \text{where} \quad f_k = kF_s/N$$

We let $K = N/2$. For K great enough, X_1, ..., X_K is a sequence of M-ary independent circularly gaussian random vectors with zero-mean and respective covariance:

$$\Gamma_k(\alpha) = \gamma_k d_k(\theta) d_k^H(\theta) + \sigma^2 I_M \tag{3}$$

where $d_k(\theta)$ is an M-ary vector whose the m-th entry writes $D_{k,m} = e^{-2j\pi f_k r_m^T \theta}$ and where $\gamma_k = \gamma_s(f_k)$.

The expression (3) can be rewritten:

$$\Gamma_k(\alpha) = \gamma_k D_k(\theta) \, C_k \, D_k^H(\theta) + \sigma^2 I_M \tag{4}$$

where $D_k(\theta)$ is an M-ary diagonal matrix whose the m-th diagonal entry writes $D_{k,m} = e^{-2j\pi f_k r_m^T \theta}$ and where $C_k = \mathbb{1}\mathbb{1}^T$ which is a rank 1 matrix. Under LOC, C_k is no more a rank 1 projector.

From (4), we get:

$$\Gamma_{k,m,m'} = \gamma_k C_{k,m,m'} e^{-2j\pi f_k (r_m - r_{m'})^T \theta} + \sigma^2 \delta_{m,m'}$$

We let $\text{SNR}_k = \gamma_k/\sigma^2$ the signal-to-noise ratio at the frequency f_k. It follows that the MSC at the frequency f_k between any two sensors writes:

$$\text{MSC}_{k,m,m'} = \frac{1}{(1 + \text{SNR}_k^{-1})^2} \, C_{k,m,m'}^2 \tag{5}$$

5

The level the MSC is due in one hand to the noise and on the other to the LOC.

A commonly used model for the LOC is for $1 \leq m \leq M$ and $1 \leq m' \leq M$:

$$C_{k,m,m'} = e^{-2\pi^2 f_k^2 (r_m - r_{m'})^T S_\theta (r_m - r_{m'})} \tag{6}$$

where S_θ is a 3×3 covariance matrix depending on 3 free parameters. Using (5), we get:

$$\log \mathrm{MSC}_{k,m,m'} = -2\log(1 + \mathrm{SNR}_k^{-1}) - 4\pi^2 f_k^2 (r_m - r_{m'})^T S_\theta (r_m - r_{m'})$$

In the simple case where $_\theta$ is proportional to the identity, and omitting the frequency index k, this expression simplifies as:

$$\log \mathrm{MSC}_{m,m'} = \beta_0 + \beta_1 d_{m,m'}$$

If we assume constant SNR in consecutive windows we can estimate β_0 and β_1.

Likelihood function

In summary we can write that:

$$(X_1, \ldots, X_K) \sim \prod_{k=1}^{K} \mathcal{N}_c(x_k; 0_M, \Gamma_k(\alpha))$$

and the likelihood writes:

$$\mathcal{L}(\alpha) = \sum_{k=1}^{K} \log \det \Gamma_k(\alpha) + \mathrm{trace}\left(\Gamma_k^{-1}(\alpha) X_k X_k^T\right) \tag{7}$$

where the parameter α consists of

$$\alpha = \{\theta_1, \theta_2, \theta_3, \gamma_1, \ldots, \gamma_K, S_\theta, \sigma^2\} \in \mathbb{R}^3 \times \mathbb{R}^{+K} \times \mathcal{M}_3^+ \times \mathbb{R}^+ \tag{8}$$

where $\mathcal{M}_3^+$ is the set of 3×3 positive matrices. The size of the parameter α is $K+7$ whereas the number of observations is $2MK$.

Another way is to characterize the slowness by azimuth, elevation and velocity and the LOC by the respective standard deviations of them, see expression (B.1).

2 Validation

Forsake of simplicity we assume that $S_\theta = \beta I_3$. Carrying this expression in (??) we get:

$$\log C_{k,\ell\ell'}(\beta) = -\beta f_k^2 d_{\ell,\ell'}^2 \tag{9}$$

meaning that the the log of MSC is proportional to the square to both frequency and distance. In the following that β is known and has been obtained on a large campaign of studies. The parameter reduces to:

$$\alpha \;\; = \;\; \{\theta_1, \theta_2, \theta_3, \sigma^2, \gamma_1, \ldots, \gamma_K\} \in \mathbb{R}^3 \times \mathbb{R}^+ \times \mathbb{R}^{+K} \tag{10}$$

In the following we have to consider the derivative of the likelihood, expression (7), w.r.t. the different components of α. For that we consider that the derivatives are taken in the order of the components of α as given by the expression (10).

To studying the relationship between the MSC and the inter-distance between sensors we need (i) a station with a large number of sensors and located in such a way that the distribution of the inter-distances is quiet uniform, (ii) a coherent source almost permanently present. A good example is the microbarom, but unfortunately it covers a small part of the frequency bandwidth of interest. We have retained

- station IS37 with microbarom,

- station IS27 with microbarom,

- station IS55 with microbarom,

- station IS31 with gas flare,

3   Cramer-Rao bound (CRB)

The (ℓ, ℓ') entry of the Fisher information matrix associated to the parameter α writes:

$$\mathrm{FIM}_{\ell,\ell'}(\alpha) \;\; = \;\; \sum_{k=1}^{K} \mathrm{trace}\left(\Gamma_k^{-1} \times \partial_\ell \Gamma_k \times \Gamma_k^{-1} \times \partial_{\ell'} \Gamma_k\right) \tag{11}$$

where $1 \le \ell, \ell' \le K+4$ and where $\partial_\ell \Gamma_k$ is the partial derivative w.r.t. α. We let:

$$\dot{d}_{k,\ell}(\theta) \;\; = \;\; \begin{bmatrix} -2j\pi f_k r_{1,\ell}\, e^{-2j\pi f_k r_1^T \theta} \\ \vdots \\ -2j\pi f_k r_{M,\ell}\, e^{-2j\pi f_k r_M^T \theta} \end{bmatrix}$$

Then for $\ell = 1, 2, 3$:

$$\begin{aligned} \partial_\ell \Gamma_k \;\; &= \;\; \gamma_k \mathrm{diag}\left(\dot{d}_{k,\ell}(\theta)\right) C_k(\beta) D_k^H(\theta) + \gamma_k D_k(\theta) C_k(\beta) \mathrm{diag}\left(\dot{d}_{k,\ell}(\theta)\right)^H \\ &= \;\; 2\,\gamma_k \mathcal{R}\left(\mathrm{diag}\left(d_{k,\ell}(\theta)\right) C_k(\beta) D_k^H(\theta)\right) \end{aligned}$$

It is worth to notice that, if $\beta \approx +\infty$, $C_k = I_M$ and $\partial_\ell \Gamma_k = 0$ which is normal because in this case Γ_k does not depend on θ. For the derivation w.r.t. σ^2 we have:

$$\partial_4 \Gamma_k \;\; = \;\; I_M$$

A direct consequence is that the FIM, w.r.t. the components (10) of α is of this shape:

$$
\text{FIM} \;=\;
\begin{bmatrix}
F_{11} & F_{12} & F_{13} & 0 & 0 & \ldots & 0 \\
F_{21} & F_{22} & F_{23} & 0 & 0 & \ldots & 0 \\
F_{31} & F_{32} & F_{33} & 0 & 0 & \ldots & 0 \\
0 & \ldots & 0 & F_{\sigma^2} & & & \\
0 & \ldots & 0 & & & & \\
\vdots & \ddots & \vdots & & & F_{\gamma,K,K} & \\
0 & \ldots & 0 & & & &
\end{bmatrix}
$$

Because the CRB on the estimation of θ is given by the 3×3 top-left matrix of the inverse of FIM, we are only concern with F_{123} by taking F_{123}^{-1}. Then the CRB w.r.t. the azimuth, elevation and velocity can be derived using the Jacobian of the one-to-one mapping between the slowness and the vector (a, e, c) where a, e and c denote respectively the azimut, the elevation and the velocity. The same calculation can be used to derive the CRB w.r.t. the azimuth and the trace velocity.

Appendix A

Transform an array in isotrope array

We consider an arbitrary array whose locations are given by:

$$X = \begin{bmatrix} x_{1,1} & \dots & x_{1,M} \\ x_{2,1} & \dots & x_{2,M} \\ x_{3,1} & \dots & x_{3,M} \end{bmatrix}$$

It follows that

$$XX^T = \sum_{i=1}^{d} \mu_i \xi_i \xi_i^T$$

where $0 \le \mu_i \le \alpha_0$. The new array writes

$$Y = \begin{bmatrix} X & \sqrt{(\alpha_0 - \mu_1)}\xi_1 & \sqrt{(\alpha_0 - \mu_2)}\xi_2 & \sqrt{(\alpha_0 - \mu_3)}\xi_3 \end{bmatrix}$$

It is easy to verify that $YY^T = \alpha_0 I_d$.

Appendix B

Coherence

The LOC is modeled by the random slowness:

$$\Theta = \theta_0 + \epsilon$$

where θ_0 is a 3D deterministic vector and ϵ a 3D random vector with zero-mean and distribution probability density denoted p_ϵ. The spectral matrix of the M-ary random vectors associated to the sensor locations has the following entry:

$$
\begin{aligned}
S_{\ell,\ell'}(f) &= \gamma(f) \int_{\mathbb{R}^3} e^{-2j\pi f(r_\ell - r_{\ell'})^T t} p_e(t - \theta_0) dt \\
&= \gamma(f) e^{-2j\pi f(r_\ell - r_{\ell'})^T \theta_0} \int_{\mathbb{R}^3} e^{-2j\pi f(r_\ell - r_{\ell'})^T t} p_e(t) dt \\
&= \gamma(f) e^{-2j\pi f(r_\ell - r_{\ell'})^T \theta_0} \times \Phi_e(2\pi f(r_\ell - r_{\ell'}))
\end{aligned}
$$

where $\Phi_e(v)$ denotes the characteristic function of the r.v. e. The deterministic case where $\epsilon = 0$ leads to:

$$S_{\ell,\ell'}(f) = \gamma(f) e^{-2j\pi f(r_\ell - r_{\ell'})^T \theta_0}$$

and therefore $S(f) = d(f)d^H(f)$ where d is a complex vector whose the ℓ-entry writes $e^{-2j\pi f r_\ell^T \theta_0}$. Hence S is a projector of rank 1 and corresponds to a pure coherent case.

In the general case:

$$S(f) = \gamma_s(f) D(f) C(f) D^H(f)$$

$D(f)$ is a diagonal matrix whose the ℓ-th diagonal entry is $e^{-2j\pi f r_\ell^T \theta_0}$ and $C(f)$ a matrix whose the (ℓ, ℓ')-entry writes:

$$C_{\ell,\ell'}(f) = \Phi_e(2\pi f(r_\ell - r_{\ell'}))$$

If ϵ is gaussian with covariance matrix Σ_θ:

$$\Phi_e(2\pi f(r_\ell - r_{\ell'})) = e^{-2\pi^2 f^2 (r_\ell - r_{\ell'})^T \Sigma_\theta (r_\ell - r_{\ell'})}$$

Σ_θ depends on 6 free parameters.

Remark 6 (Unities) *the covariance matrix Σ_θ is in s^2/m^2. It follows that $f^2\Sigma_\theta$ is homogeneous at the inverse of the wavelength square.*

Using (C.1), we can derive an expression in term of azimuth a, elevation e and velocity c. We have:

$$\Sigma_\theta = J\,\Sigma_{\mathrm{aec}}\,J^T \tag{B.1}$$

Then

$$
\begin{aligned}
\log \Phi_e(2\pi f(r_\ell - r_{\ell'})) &= -2\pi^2 \frac{f^2}{c^2}(r_\ell - r_{\ell'})^T K(a,e,c)\Sigma_{\mathrm{aec}}K^T(a,e,c)(r_\ell - r_{\ell'}) \\
&= -2\pi^2(\rho_\ell - \rho_{\ell'})^T K_{\mathrm{aec}}\Sigma_{\mathrm{aec}}K^T_{\mathrm{aec}}(\rho_\ell - \rho_{\ell'})
\end{aligned}
$$

where $\rho = r/\lambda$ with $\lambda = c/f$ and

$$
K_{\mathrm{aec}} = \begin{bmatrix}
\cos(a)\cos(e) & \sin(a)\sin(e) & c^{-1}\sin(a)\cos(e) \\
-\sin(a)\cos(e) & -\cos(a)\sin(e) & -c^{-1}\cos(a)\cos(e) \\
0 & \cos(e) & -c^{-1}\sin(e)
\end{bmatrix}
$$

A simple case is if we take Σ_{aec} diagonal and hence depending on only 3 free parameters.

Appendix C

One-to-one mappings and Jacobians

θ **to** (a, e, c)

If we consider the one-to-one mapping θ to (a, e, c) in $(0, 2\pi) \times (-\pi/2, \pi/2) \times \mathbb{R}^+$, we can write:

$$\begin{cases} \theta_1 = -c^{-1}\sin(a)\cos(e) & \theta_1 \in \mathbb{R} \\ \theta_2 = c^{-1}\cos(a)\cos(e) & \theta_2 \in \mathbb{R} \\ \theta_3 = c^{-1}\sin(e) & \theta_3 \in \mathbb{R} \end{cases} \quad \begin{cases} a = \arg(\theta_2 - j\theta_1) & a \in (0, 2\pi) \\ e = \arg\sin(c\theta_3) & e \in (-\pi/2, \pi/2) \\ c = (\theta_1^2 + \theta_2^2 + \theta_3^2)^{-1/2} & c \geq 0 \end{cases}$$

whose the Jacobian is

$$\begin{aligned} J(a, e, c) &= \begin{bmatrix} -c^{-1}\cos(a)\cos(e) & c^{-1}\sin(a)\sin(e) & c^{-2}\sin(a)\cos(e) \\ -c^{-1}\sin(a)\cos(e) & -c^{-1}\cos(a)\sin(e) & -c^{-2}\cos(a)\cos(e) \\ 0 & c^{-1}\cos(e) & -c^{-2}\sin(e) \end{bmatrix} \quad (\text{C.1}) \\ &= c^{-1}\begin{bmatrix} \cos(a)\cos(e) & \sin(a)\sin(e) & c^{-1}\sin(a)\cos(e) \\ -\sin(a)\cos(e) & -\cos(a)\sin(e) & -c^{-1}\cos(a)\cos(e) \\ 0 & \cos(e) & -c^{-1}\sin(e) \end{bmatrix} \\ &= c^{-1}K_{\text{aec}} \end{aligned}$$

θ **to** (a, e, v)

If we consider the one-to-one mapping θ to (a, e, v) in $(0, 2\pi) \times (-\pi/2, \pi/2) \times \mathbb{R}^+$, we can write:

$$\begin{cases} \theta_1 = -v^{-1}\sin(a) & \theta_1 \in \mathbb{R} \\ \theta_2 = v^{-1}\cos(a) & \theta_2 \in \mathbb{R} \\ \theta_3 = v^{-1}\tan(e) & \theta_3 \in \mathbb{R} \end{cases} \quad \begin{cases} a = \arg(\theta_2 - j\theta_1) & a \in (0, 2\pi) \\ e = \arg\tan(v\theta_3) & e \in (-\pi/2, \pi/2) \\ v = (\theta_1^2 + \theta_2^2)^{-1/2} & c \geq 0 \end{cases}$$

whose the Jacobian is

$$J(a, e, v) = \begin{bmatrix} -v^{-1}\cos(a) & 0 & v^{-2}\sin(a) \\ -v^{-1}\sin(a) & 0 & -v^{-2}\cos(a) \\ 0 & v^{-1}/\cos^2(e) & -v^{-2}\tan(e) \end{bmatrix} \quad (\text{C.2})$$

Appendix D

Estimation of parameters

we have

$$
\begin{aligned}
\log C_{k,m,m'} &= -2\pi^2 f_k^2 (r_m - r_{m'})^T \Gamma (r_m - r_{m'}) \\
&= -2\pi^2 f_k^2 (r_{m,1} - r_{m',1})^2 \Gamma_{11} - 2\pi^2 f_k^2 (r_{m,2} - r_{m',2})^2 \Gamma_{22} - 4\pi^2 f_k^2 (r_{m,1} - r_{m',1})(r_{m,2} - r_{m',2}) \Gamma_{12} \\
&= \mu_1(k,m,m') \Gamma_{11} + \mu_2(k,m,m') \Gamma_{22} + 2\mu_3(k,m,m') \Gamma_{12}
\end{aligned}
$$

Enumerating $k = 1 : K$, $m = 1 : M$, $m' = 1 : M$ with $m' > m$ leads to $KM(M-1)/2$ equations with 3 unknowns.

Appendix E

inverse of Γ_k

Here we give the analytical expression of the inverse of $\Gamma_k(\alpha) = \gamma_k D_k(\theta)\, C_k\, D_k^H(\theta) + \sigma^2 I_M$

$$
\begin{aligned}
\Gamma_k^{-1}(\alpha) &= \sigma^{-2} I_M - \sigma^{-4} D_k(\theta)(C_k^{-1} + \sigma^{-2} D_k^H(\theta) D_k(\theta))^{-1} D_k^H(\theta) \\
&= \sigma^{-2} I_M - \sigma^{-4} D_k(\theta)(C_k^{-1} + \sigma^{-2} I)^{-1} D_k^H(\theta)
\end{aligned}
$$