2

LOW-RANK APPROXIMATIONS FOR FIELDS

TRANSFORMATIONS

In this chapter we present the approximation techniques that have been applied on the scalar near fields and on the scalar near-to-far process, in order to find suitable spanning spaces for the beamsteering operation. We recall that what we are seeking for is an adequate approximation method, with controllable accuracy, in order to apply it to a full FEM model of a phased array.

We will first begin with a *Singular Value Decomposition* (SVD) based technique _a number of_ that will be applied on the near fields in order to define the possible order of a re- _outputs_ duced model. Then a *polynomial interpolation* and a *Discrete Fourier Transform* (DFT) based technique, the *DFT-truncation*, will be tested on the near-to-far _field_ operator, seeking for a reduced basis to express the N2F process. The accuracy of these techniques will be given in terms of error induced to the pattern, relatively to the non approximated N2F pattern and on the directly computed pattern.

47

2.1 Approximations of the near fields

To introduce the SVD, we first discuss on the *Proper Orthogonal Decomposition* (POD), for which the SVD is the discrete case. The POD is a powerful and elegant method of data analysis aimed of obtaining low-dimensional approximate descriptions of high dimensional processes. The POD is often used to extract *mode shapes* or basis functions, from experimental data or detailed simulations of high-dimensional systems for subsequent use in *Petrov-Galerkin projections* [13] that yield low dimensional dynamical models [14].

2.1.1 Proper orthogonal decomposition

Let $z(r) \in \mathbb{C}$ be a function of the position vector $r \in \mathbb{R}^3$ over the domain of interest $S \subseteq \mathbb{R}^3$. The chosen variable r and the domain S, particularly related to the electromagnetic problem we wish to approximate, may be straightforwardly extended to any variables (even more than one assuming separability of the variables) and domains for a more general formulation. Now, suppose we wish to approximate the function $z(r)$ as a finite sum

$$z(r) = \sum_{m=1}^{M} a_m \phi_m(r), \qquad a_m, \; \phi_m(r) \in \mathbb{C}, \qquad (2.1.1)$$

with the reasonable expectation that the approximation becomes exact in the limit as $M \to \infty$, except possibly on a set of measure zero [15]. The representation (2.1.1) of the function $z(r)$ is not unique. If the domain if r is a bounded interval S, then the functions $\phi_m(r)$ can be chosen as a Fourier series, or Legendre polynomials, or Chebyshev polynomials, and so on. For each such choice of a sequence $\phi_m(r)$ that forms a basis for some suitable class of functions $z(r)$, assuming these functions bounded and integrable, the sequence of coefficients[1] a_m is different. The POD is concerned with one possible choice of the functions $\phi_m(r)$. If we have chosen or-

---

[1] We have supposed our function $z(r)$ to be time independent, as we are interested in the approximation of fields in a frequency domain formulation. However, if we had $z(r,t)$ with t the time variable, the coefficients a_m would become functions $a_m(t)$ to hold the time dependence.

thonormal basis functions, i.e.

$$\oint_S \phi_m^\dagger(\boldsymbol{r}')\phi_n(\boldsymbol{r}')\,dS' = \delta_{mn}, \tag{2.1.2}$$

where $\delta_{mn} = 1$ if $m = n$, 0 otherwise is the *Kronecker's delta*, then

$$a_m = \oint_S \phi_m^\dagger(\boldsymbol{r}')z(\boldsymbol{r}')\,dS', \tag{2.1.3}$$

that is the coefficient a_m depends only on the related function $\phi_m(\boldsymbol{r}')$ and not on the others. By the integral (2.1.3), the function $z(\boldsymbol{r}')$ is said to be *tested* by the function $\phi_m(\boldsymbol{r}')$, operation after which we derive a *weighting* coefficient a_m of $\phi_m(\boldsymbol{r}')$ for the approximating sum 2.1.1. If the orthonormal functions are ordered such that the best approximant $\phi_m(\boldsymbol{r})$ of $z(\boldsymbol{r})$ have lower indices, that is, as M grows $\to \infty$, the series (2.1.1) converges monotonically to $z(\boldsymbol{r})$ in a *least squares sense* ($\mathbb{L}_2$-norm error), (2.1.1) is then called the POD of $z(\boldsymbol{r})$.

↳ You described Galerkin projection. POD deals with the choice of $\phi_m(r)$

2.1.2 Singular value decomposition

As said before, the SVD is the discrete case of the POD. We proceed with the definition of the SVD. [19] [20]

Definition 1 *Consider* $\mathbf{A} \in \mathbb{C}^{m \times n}$, $m \geq n$. *The triple factorization* $\mathbf{A} =: \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H$, $\mathbf{V}^H$ *being the conjugate transpose* $\mathbf{V}$, *is called singular value decomposition (in the reduced form) of* $\mathbf{A}$ *if* $\mathbf{U} \in \mathbb{C}^{m \times n}$, $\mathbf{V} \in \mathbb{C}^{n \times n}$, $\boldsymbol{\Sigma} \in \mathbb{R}^{n \times n}$ *are such that* $\mathbf{U}^H\mathbf{U} = \mathbf{V}^H\mathbf{V} = \mathbf{V}\mathbf{V}^H = \mathbf{I}_n$, $\mathbf{I}_n$ *being the identity matrix, and* $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$.

The matrix $\mathbf{U}$ *consists of* n *orthonormal eigenvectors associated with the* n *largest eigenvalues of* $\mathbf{A}\mathbf{A}^H$, *and the matrix* $\mathbf{V}$ *consists of the orthonormal eigenvectors of* $\mathbf{A}^H\mathbf{A}$. *The columns of* $\mathbf{U}$ *are called the left singular vectors, the columns of* $\mathbf{V}$ *are called the right singular vectors. The diagonal elements of* $\boldsymbol{\Sigma}$ *are the nonnegative square roots of the eigenvalues of* $\mathbf{A}^H\mathbf{A}$ *and they are called singular values. We shall assume that* $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$. *Thus if* $\mathrm{rank}(\mathbf{A}) = r$, $\sigma_{r+1} = \sigma_{r+2} = \ldots = \sigma_n = 0$. *In other terms,*

$$\mathbf{U} = [\mathbf{U}_1\,\mathbf{U}_2], \qquad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_1 & 0 \\ 0 & \boldsymbol{\Sigma}_2 \end{pmatrix} \in \mathbb{R}^{n \times n}, \qquad \mathbf{V} = [\mathbf{V}_1\,\mathbf{V}_2]$$

$$\Sigma_1 = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix}, \qquad \Sigma_2 = 0 \in \mathbb{R}^{n-r \times n-r}$$

$$\mathbf{U}_1 \in \mathbb{C}^{m \times r}, \mathbf{U}_2 = 0 \in \mathbb{C}^{m \times n-r}, \qquad \mathbf{V}_1 \in \mathbb{C}^{r \times n}, \mathbf{V}_2 = 0 \in \mathbb{C}^{n-r \times n}$$

As a consequence of definition 1, the following relations on four fundamental spaces of **A** hold

$$
\begin{aligned}
\mathrm{colsp}\{\mathbf{A}\} &= \mathrm{colsp}\{\mathbf{U}_1\}, \\
\mathrm{colsp}\{\mathbf{A}^H\} &= \mathrm{colsp}\{\mathbf{V}_1\}, \\
\ker\{\mathbf{A}\} &= \mathrm{colsp}\{\mathbf{U}_2\}, \\
\ker\{\mathbf{A}^H\} &= \mathrm{colsp}\{\mathbf{V}_2\} \oplus \; \mathcal{U}^{\perp} \leftarrow \textit{for } m \neq n
\end{aligned}
\tag{2.1.4}
$$

where *colsp* is the space spanned by the columns of a matrix, that is the *range*, and *ker* is the *kernel* or *null space* of a matrix. [19]

By analogy with the POD previously presented, the left singular vectors **U** correspond to the orthonormal basis functions $\phi_m(\mathbf{r})$. These singular vectors actually constitute a space of *testing vectors* for matrices spanned by the same range of (A) bold that provides the coefficients of an expansion in terms of vectors **U** (the *trial* vectors which in this case are the same as the testing vectors). The coefficients a_m can be viewed as the testing matrix multiplication $\mathbf{U}^H \mathbf{A} = \mathbf{K}$, $\mathbf{K} \in \mathbb{C}^{m \times n}$ being a matrix of coefficients that weights the basis **U**. In other terms, if rank$(\mathbf{A}) = r$ we can build a matrix $\mathbf{\Pi} = \mathbf{U}\mathbf{U}^H \in \mathbb{C}^{m \times m}$, an *orthogonal projection*[2] of any matrix onto the r dimensional subspace spanned by the columns of **U** (colsp$\{\mathbf{A}\}$) and along the kernel of $\mathbf{U}^H$ (left singular vectors associated to null singular values, that i ker$\{\mathbf{A}\}$).

Another interesting issue of the SVD is the fact that the singular vectors are ordered in function of their importance in the representation of the matrix **A**, the singular values being ordered monotonically in a decreasing order. We can thus select $q \leq r$ left singular vectors, associated to the first q singular values, to build a projector

---

[2] A linear mapping that provides a change of basis. In this case, we change the basis into an orthonormal one.

↳ Π is not really a change of basis. Multiplication with $\mathcal{U}^H$ is change of basis.

$\mathbf{\Pi}_q \in \mathbb{C}^{m \times m}$ and $\text{rank}\left(\mathbf{\Pi}_q\right) = q$ such that $\hat{\mathbf{A}}_q = \mathbf{\Pi}_q \mathbf{A}$ is a qth order *low-rank approximation* of the matrix $\mathbf{A}$, achieved by a projection technique. By analogy with the POD, q corresponds to the degree of approximation M. Furthermore, the following important theorem on the approximation holds [19]

Theorem 1 *Let* $\mathbf{A} \in \mathbb{C}^{m \times n}$, $m \geq n$ *be a matrix with singular value decomposition* $\mathbf{U \Sigma V}^H$. *For any* $q : \quad 0 \leq q \leq r$, $\mathbf{A}_q = \sum_{j=1}^r \sigma_j \mathbf{u}_j \mathbf{v}_j^\dagger$, $\mathbf{u}_j, \mathbf{v}_j$ *column vectors. If* $q = \min(m, n) = \text{rank}(\mathbf{A}) = r, \sigma_{q+1} = 0$. *Then*

$$\|\mathbf{A} - \mathbf{A}_q\|_2 = \inf_{\substack{\mathbf{B} \in \mathbb{C}^{m \times n} \\ \text{rank}(\mathbf{B}) \leq q}} \|\mathbf{A} - \mathbf{B}\|_2 = \sigma_{q+1}.$$

Theorem 1 states that we can compute the error for the approximation and that the approximation is optimal with respect to the spectral norm[3].

2.1.3   Space spanned by the scanned near fields

Along a beamsteering process, the dimension of the *scan angle space* that can hold all the configurations of the near fields is *a priori* unknown. The SVD comes out helping us to find it. We may first collect some *pictures* of the near fields behavior for some scan angles (θ_x, θ_y) (section 1.2.2.2) to build our space. In other words, the near fields sampled values for each n scan angle chosen will constitute spanning vectors ψ_n such that $\mathbf{A}_\theta = \text{span}\left\{\bigcup_{n=1}^N \psi_n\right\}$. The rank will be checked performing an SVD on $\mathbf{A}_\theta = \mathbf{U}_\theta \mathbf{\Sigma}_\theta \mathbf{V}_\theta^H$ and looking at the singular values to derive informations on the importance of the column vectors collected. Relying on definition 1, the singular values of $\mathbf{A}_\theta$ become null as soon as the number of spanning vectors N supersedes the rank of the scan angle space, expressing in such a way its dimension.

The choice of the scan angles (θ_x, θ_y) is arbitrary, and we also don't know *a priori* which sequence may lead to a rapid convergence into the full rank scan angle space. In a first attempt, we choose the broadside direction, obtaining the vector $\psi_{(0°,0°)}$. Then, limiting the scanning operation in the x direction, we select the end

---

[3] Natural norm induced by the $\mathbb{L}_2$-norm : $\|\mathbf{A}\|_2 = \sup_{|\mathbf{x}|_2 \neq 0} \frac{\|\mathbf{Ax}\|_2}{\|\mathbf{x}\|_2}$.

fire steering angle in that direction. Thus, we have the new spanning vector $\psi_{(90°,0°)}$. The next vectors will be chosen in between of $(0°,0°)$ and $(90°,0°)$, halving iteratively the steering intervals inserting a new vector in the middle. We obtain the sequence $\psi_{(0°,0°)}, \psi_{(90°,0°)}, \psi_{(45°,0°)}, \psi_{(22.5°,0°)}, \psi_{(67.5°,0°)}, \psi_{(11.25°,0°)}, \ldots$ and so on, until we achieve the sought full rank space. Notice this is an exponential growth of the spanning space of the kind $\dim(\mathbf{A}_\theta) = \lfloor 2^{n-1} \rfloor + 1$, $n = 0, 1, 2, \ldots$, and we will call this sequence *exponential scan angle selection*.

As we are building the space increasing progressively its dimension, we may test the basis completeness applying the projector resulting from the left singular vectors onto a picture of the near fields that is not contained in the vectors that spans $\mathbf{A}_\theta$ (otherwise we have exact approximation). We have arbitrarily chosen the scan angle $(23.3251°, 0°)$, and the fields are produced on an enclosing sphere by an array of 17 by 1 point sources (as we are scanning along θ_x, a linear array can be used to achieve faster computations), varying the spacing between the point sources.
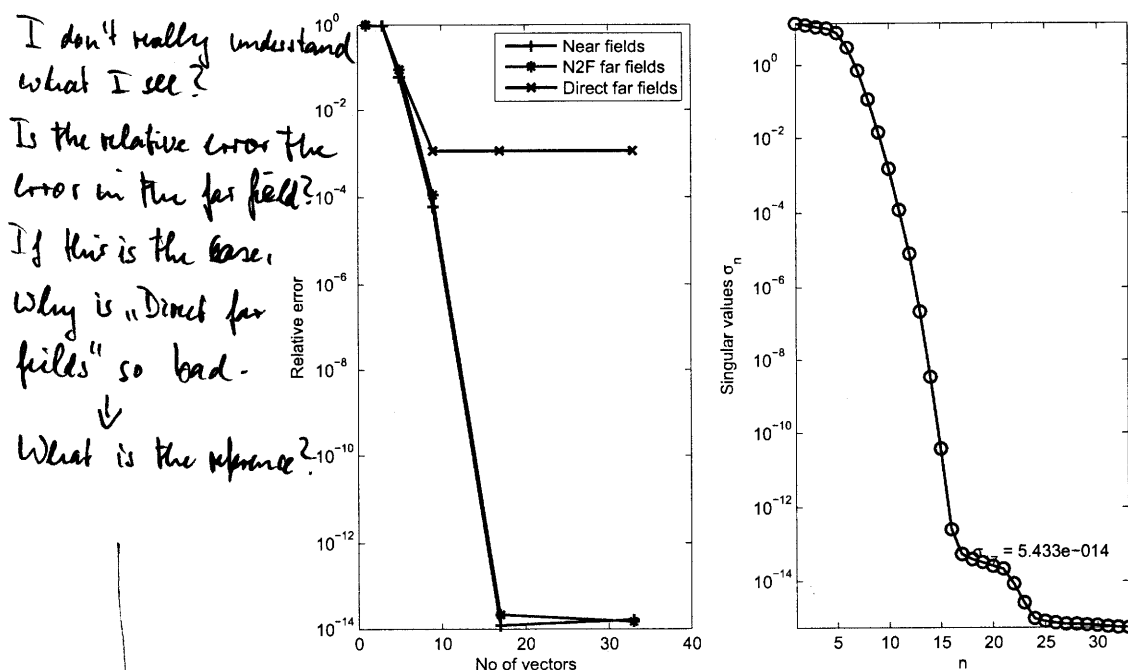
Figure 2.1: Error plot relative to the number of vectors $\psi_{(\theta_x°,0°)}$ used in the scan angle space (left) and singular values related to the scan angle space build with 33 vectors (right) for the chosen array of 17 by 1 point sources equally spaced of $\lambda/4$.
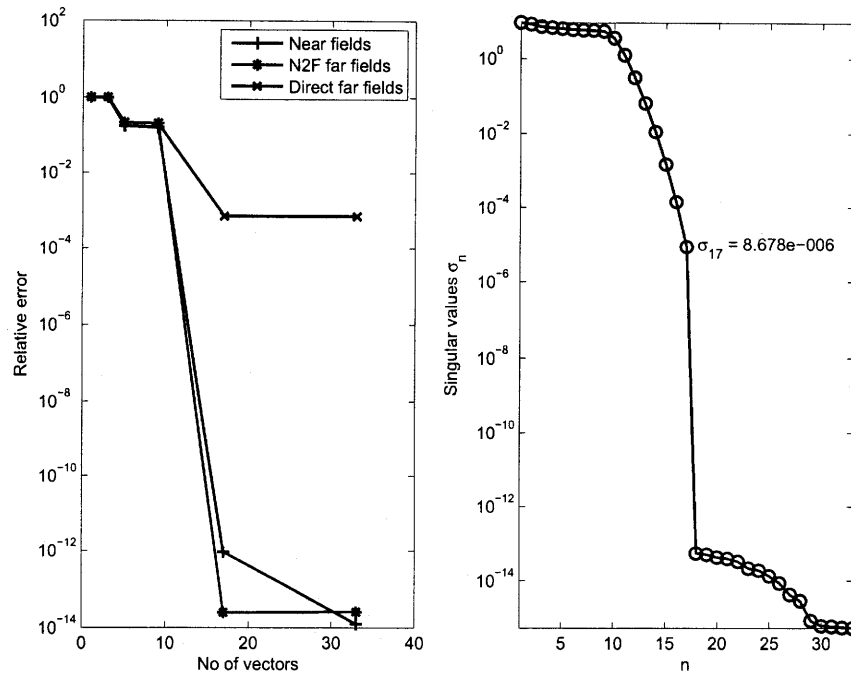
Figure 2.2: Error plot for the chosen array of 17 by 1 point sources equally spaced of $\lambda/2$.
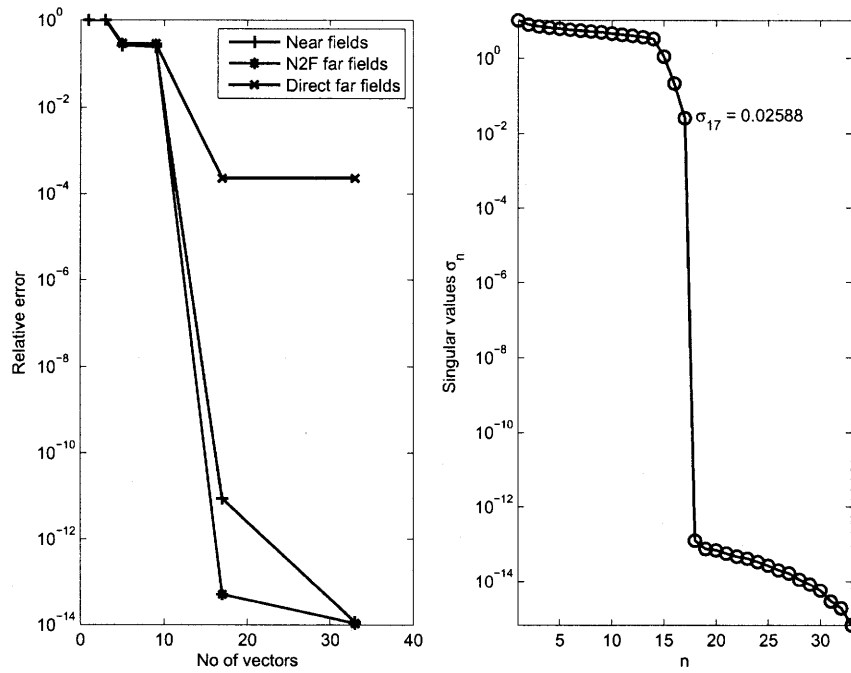

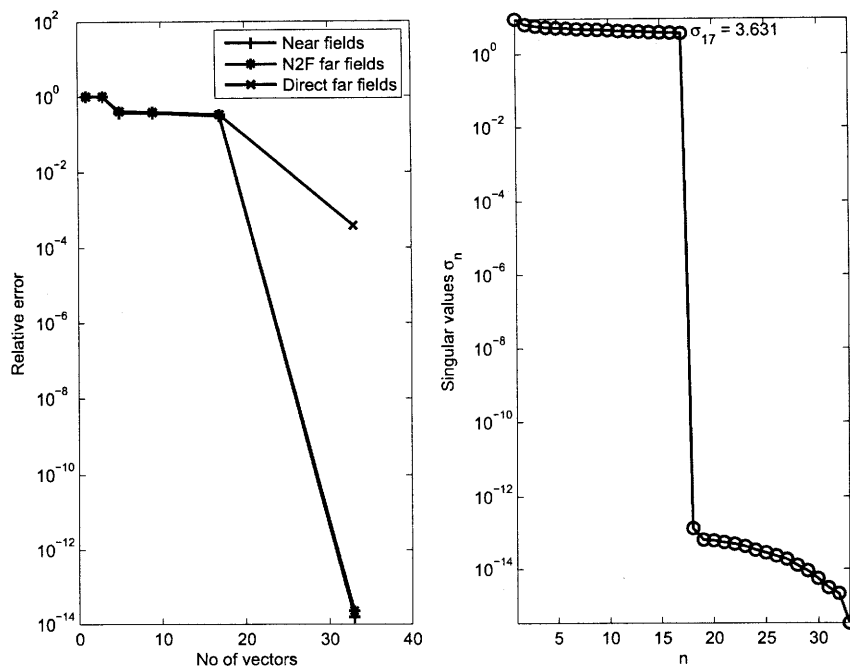Figure 2.3: Error plot for the chosen array of 17 by 1 point sources equally spaced of $3\lambda/4$.

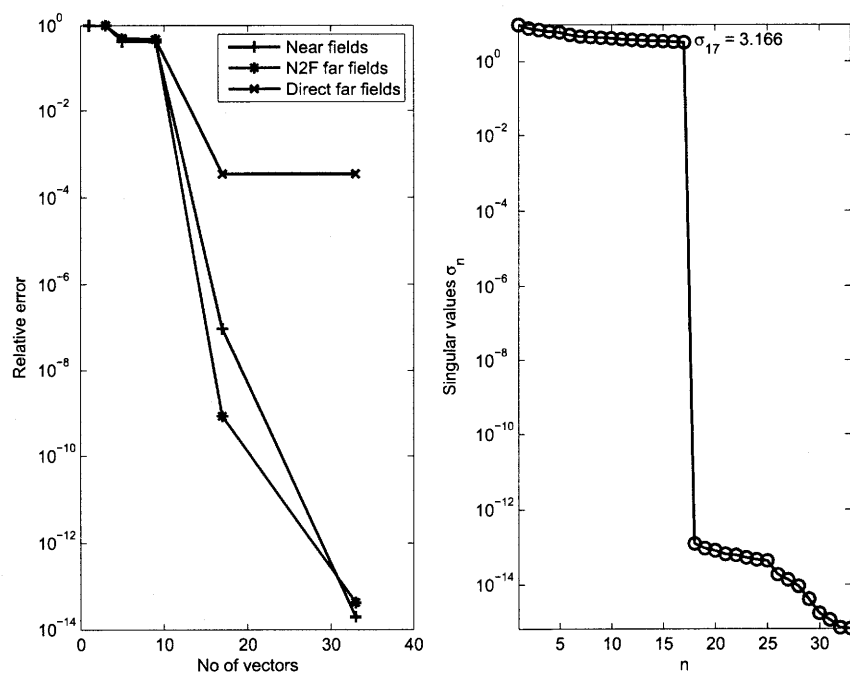Figure 2.4: Error plot for the chosen array of 17 by 1 point sources equally spaced of λ.

Figure 2.5: Error plot for the chosen array of 17 by 1 point sources equally spaced of $5\lambda/4$.

Figures 2.1-2.5 shows the error on the near fields, N2F far fields and direct far
fields while increasing the dimension of the spanning space. The limit in the spacing
between the point sources, in order to avoid the insurgence of *grating lobes* [2], that
is constructive directions of the far fields different from the desired ones, is of $\approx 0.6\,\lambda$.
However, we have extended that range to illustrate the changes in the singular values.
The direct far fields error is inferiorly limited by the sampling resolution, chosen to
be of $\lambda/10$. We can see that the singular values drop below two orders of magnitude
over the double precision used ($\approx 10^{-16}$), and we will refer to this low level as the
numerical null threshold [20], after the 17th singular value, that is the number of point
sources disposed along x. However, the relative error in the pattern is not always null
when spanning with 17 vectors. Thus, relying on definition 1, we conclude that:

Conclusion 1 *The numerical rank (the rank referred to the chosen numerical null
threshold) of the scan angle space, resulting from a beamsteering operation along a
single direction, equals the number of radiating elements along that direction. Fur-
thermore, the error in the pattern computation depends on the relative spacing be-
tween the radiating elements, and within the range that avoids grating lobes, the error
grows as the array becomes electrically larger.*

Figure 2.6 illustrates the state of the space spanned by 17 vectors for a linear array
of 17 point sources equally spaced of $\lambda/2$. We notice differences between the singular
values related to the space spanned by 17 vectors and the one the with 33 vectors
(figure 2.2), even if $\mathbf{A}_\theta\big|_{17} = \mathrm{span}\left\{\bigcup_{n=1}^{17}\psi_n\right\} \subset \mathbf{A}_\theta\big|_{33} = \mathrm{span}\left\{\bigcup_{n=1}^{33}\psi_n\right\}$. To see
how different are the basis on which the near fields picture selected is projected, let us
compare the performances of a low rank projector of dimension q, assuming $r = 17$
as the rank of the scan angle space, when the space built has 17 vectors, 33 vectors
and 65 vectors. Table 2.1 show values of relative errors in the near fields and in their
N2F patterns. The order is slightly the same for all the cases. However, the space
spanned by 17 vectors may allow an truncation order up to 12 vectors, where the N2F
error is below our accuracy criterion ($5\cdot10^{-3}$, see section 1.2.2.2) , while the others
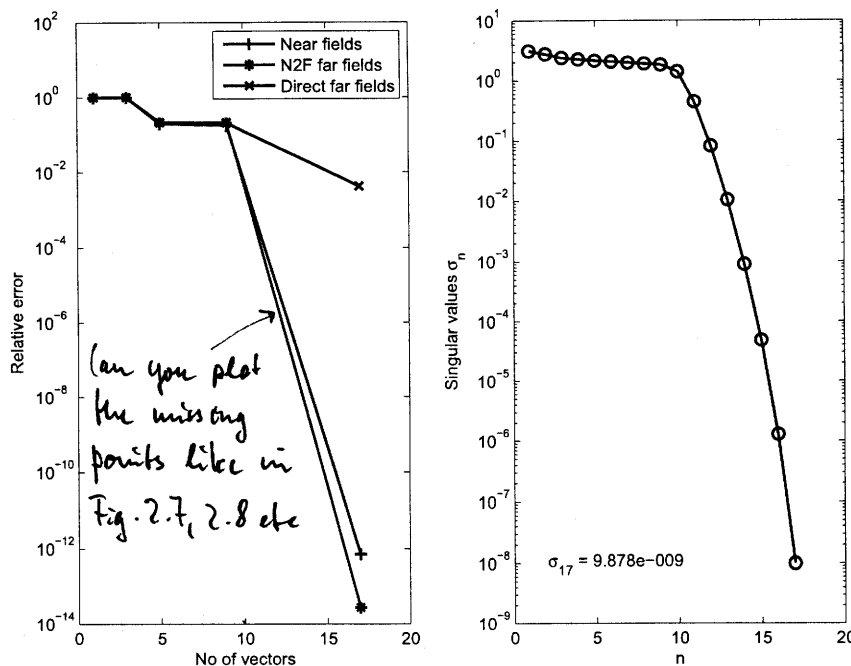would be limited to 13 vectors for the same criterion.

Figure 2.6: Error plot relative to the number of vectors $\psi_{(\theta_x°,0°)}$ up to 17 (left) and singular values related to the scan angle space build with 17 vectors (right) for an array of 17 by 1 point sources equally spaced of $\lambda/2$.

[handwritten: Are the first 17 vectors of these 33 vectors the same as the 17 vectors in the left column? If this is true, you can forget this remark.]

	17 vectors		33 vectors		65 vectors	
q	n.f. error	f.f. error	n.f. error	f.f. error	n.f. error	f.f. error
17	$1.963\ 10^{-12}$	$3.005\ 10^{-13}$	$7.153\ 10^{-14}$	$2.948\ 10^{-13}$	$6.933\ 10^{-14}$	$3.062\ 10^{-13}$
16	$2.975\ 10^{-06}$	$9.877\ 10^{-06}$	$1.233\ 10^{-07}$	$4.309\ 10^{-07}$	$1.536\ 10^{-06}$	$5.406\ 10^{-06}$
15	$2.301\ 10^{-05}$	$7.737\ 10^{-05}$	$3.008\ 10^{-05}$	0.0001065	$2.831\ 10^{-05}$	0.0001012
14	0.0001395	0.0004772	0.0002284	0.0007944	0.0002965	0.001039
13	0.0007067	0.00248	0.001282	0.004903	0.001119	0.004345
12	0.001066	0.00343	0.013	0.04836	0.01246	0.0462
11	0.03508	0.1508	0.01581	0.07227	0.01728	0.07936
10	0.2715	1.104	0.2271	0.9225	0.2049	0.8253

Table 2.1: Error in the near fields (n.f.) approximation of order q and resulting N2F far fields (f.f.).

[handwritten: I think you should mention in the text that these numbers strongly depend on the sequence of the angles chosen. Is this true? → Yes, see Fig 2.7, 2.8.]

Conclusion 2 *Reduction of the order of the near fields scan angle space can be applied, depending on the importance of the basis vectors selected, that is on the values of the singular values associated to the left singular vectors neglected in the order reduction. Furthermore, the order of the approximation depends on the spanning space dimension chosen.*

[handwritten: What are you trying to say, and how is this justified by the numerical example? It seems that increasing the space dimension brings no improvement.]
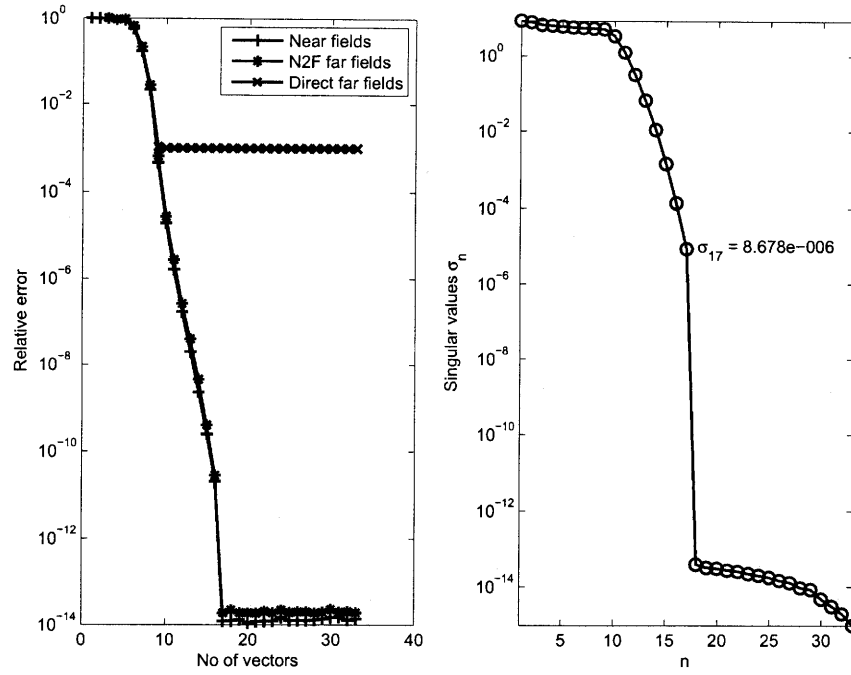
Figure 2.7: Error on the tested angle of 23.3251° and singular values for 33 vectors selected linearly with $\theta_x = 0° \to 90°$ (17 point sources linear array, spacing of $\lambda/2$).
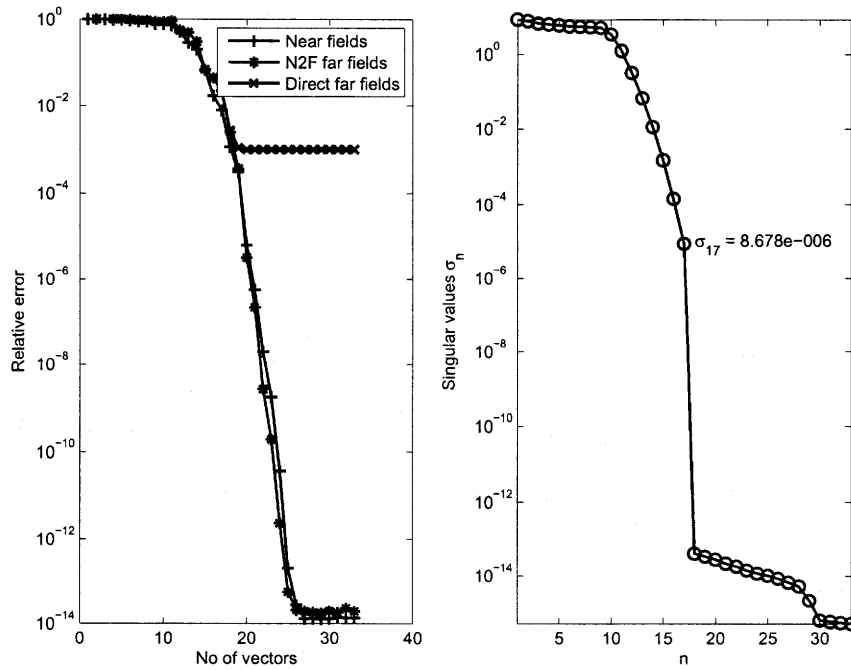

Figure 2.8: Error on the tested angle of 23.3251° and singular values for 33 vectors selected linearly with $\theta_x = 90° \to 0°$ (17 point sources linear array, spacing of $\lambda/2$).

To see whether the choice of the spanning angles is important while building $\mathbf{A}_\theta$, let us try two other sequences: the linear sequences from $\theta_x = 0° \to 90°$ and *vice versa*, with a fixed $\Delta\theta_x$. We have chosen up to 33 vectors ($\Delta\theta_x = 2.8125°$) and the resulting errors are shown in figures 2.7 and 2.8. In the first case, the tested angle of 23.3251° become included in the scan angle space interval $[0°, 25.3125°]$ with the 10th vector ($\psi_{(25.3125°,0°)}$), but the error begins its decreasing slope 5 vectors before, and the error becomes numerically null at the 17th vector (size of the array). In the second case, we need to wait until the 25th vector ($\psi_{(22.5°,0°)}$) to see the error dropping below the numerical noise floor ($< 10^{-13}$). The final 33 vectors, being the same in the two cases, are such that, obviously, we achieve the same singular values. Looking at figures 2.9
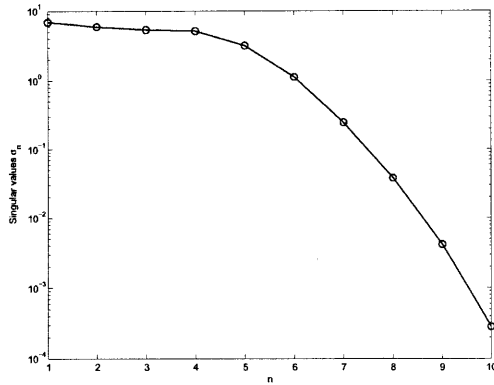

Figure 2.9: Singular values of the space spanned by 10 vectors in a linear selection of angles from $0° \to 90°$.
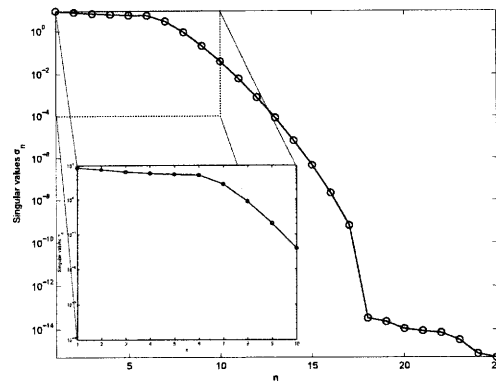
Figure 2.10: Singular values of the space spanned by 25 vectors in a linear selection of angles from $90° \to 0°$.

and 2.10, respectively, the singular values for the space spanned by only 10 vectors with the first angles selection and those for the space spanned by 25 vectors with the inverse angles selection, we notice that the slope in the singular values begins slightly earlier in the first (about 2 vectors before). This confirms that

Conclusion 3 *The selected angles for the near fields used in the scan angle space influences the convergence to the full space. Moreover, the projection of a near fields picture on that space leads to high error values if the scan angle of the projected fields is not included in the range of the selected scan angles to span the space, independently on whether the number of column vectors has superseded the rank of the space.*
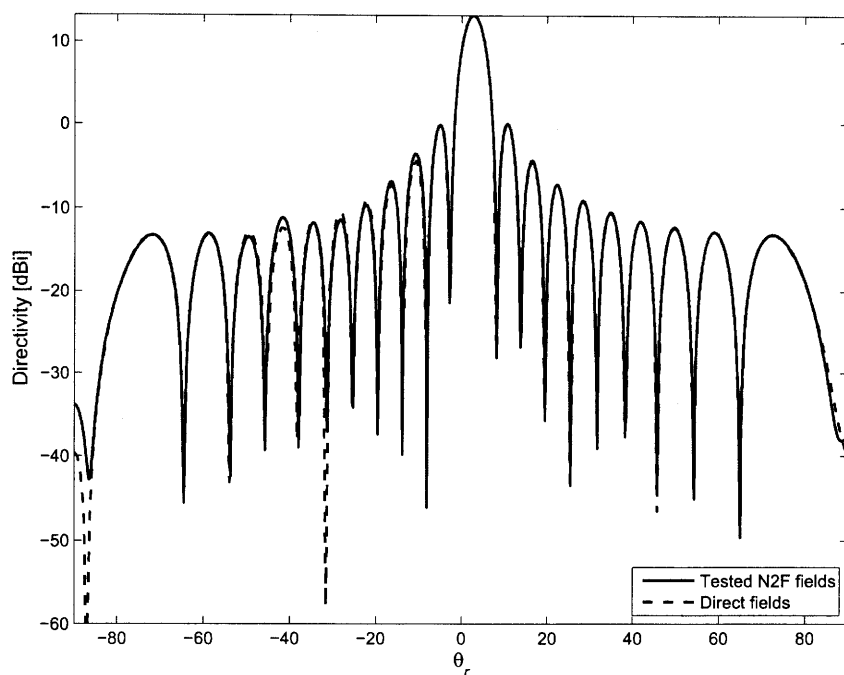
Figure 2.13: Pattern of the fields produced by a linear array of 21 point sources steered at 2.8125° and tested by a 17 vectors basis.
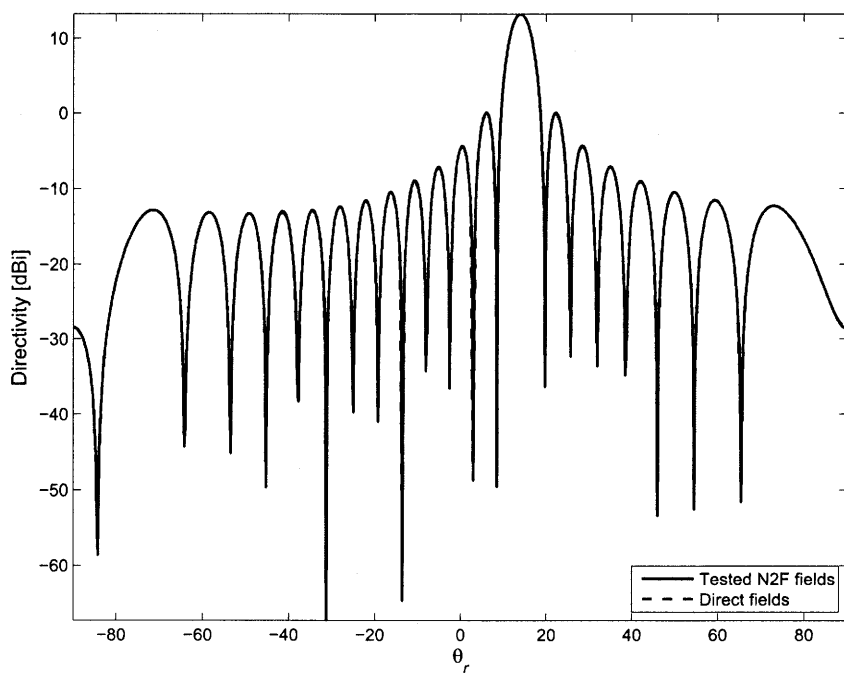

Figure 2.14: Pattern of the fields produced by a linear array of 21 point sources steered at 14.0625° and tested by a 17 vectors basis.

Let us now check how the error grows while slightly increasing the dimension of the array above 17 point sources and projecting the chosen test vector onto a space spanned by 17 vectors obtained by an exponential scan angle selection. In order to have the maximum error, the tested angle have been changed to 2.8125°, that is an angle in the middle of 0° and 5.625° for which we have collected the fields to span our scan angle space. The results are shown in figures 2.11 and 2.12.
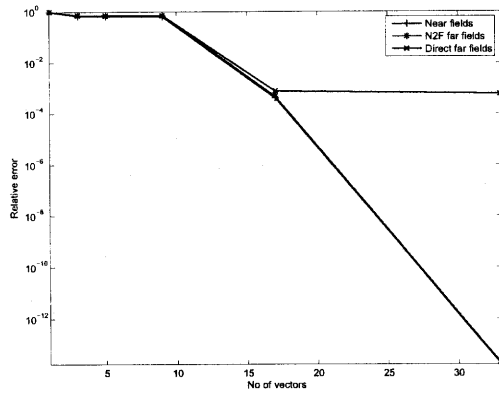

Figure 2.11: Error plot for an array of 18 point sources, spacing of $\lambda/2$.
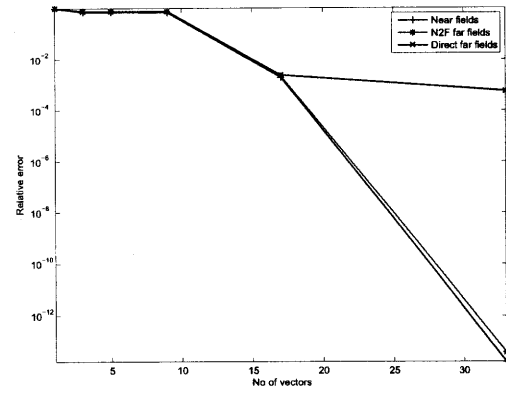
Figure 2.12: Error plot for an array of 19 point sources, spacing of $\lambda/2$.

The number of vectors that can be used as a basis for the fields produced by an array of 18 point sources, for the accuracy chosen of $5 \cdot 10^{-3}$, can be of 17. This is true for any test angle in the range $[0°, 90°]$. In effect, several angles have been tested, remaining in the middle of two spanning angles for the basis. The pattern resulting from a linear array of 21 point sources, steered at $\theta_x = 2.8125°$ and projected onto a basis of 17 vectors built by an exponential scan angle selection is shown in figure 2.13. As the N2F error, being of 0.01984 (n.f. error = 0.01724), has superseded our accuracy threshold, we can see some misfits on the lateral lobes. However, for practical applications, this can be acceptable. Now testing the pattern steered at $5 \cdot 2.8125° = 14.0625°$ (figure 2.14), we obtain a N2F error of $1.55 \cdot 10^{-3}$ (n.f. error = $1.342 \cdot 10^{-3}$). Then for $11 \cdot 2.8125° = 30.9375°$ (figure 2.15) and for $21 \cdot 2.8125° = 59.0625°$ (figure 2.16) we have, respectively, the errors $5.48 \cdot 10^{-5}$ (n.f. error = $5.042 \cdot 10^{-5}$) and $1.92 \cdot 10^{-7}$ (n.f. error = $2.82 \cdot 10^{-7}$).

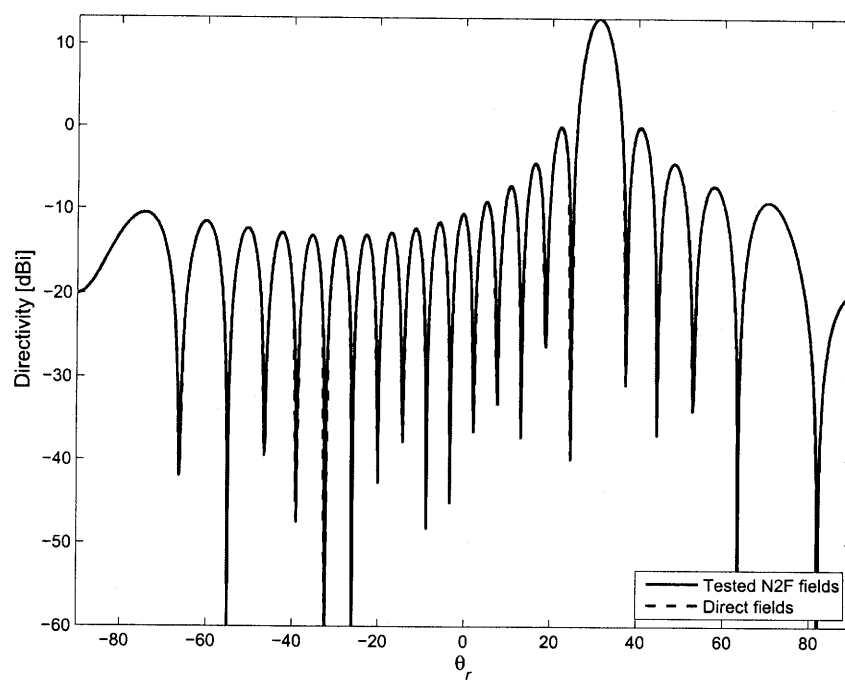We can conclude that, as the near fields approximations are more accurate the

Figure 2.15: Pattern of the fields produced by a linear array of 21 point sources steered at 30.9375° and tested by a 17 vectors basis.
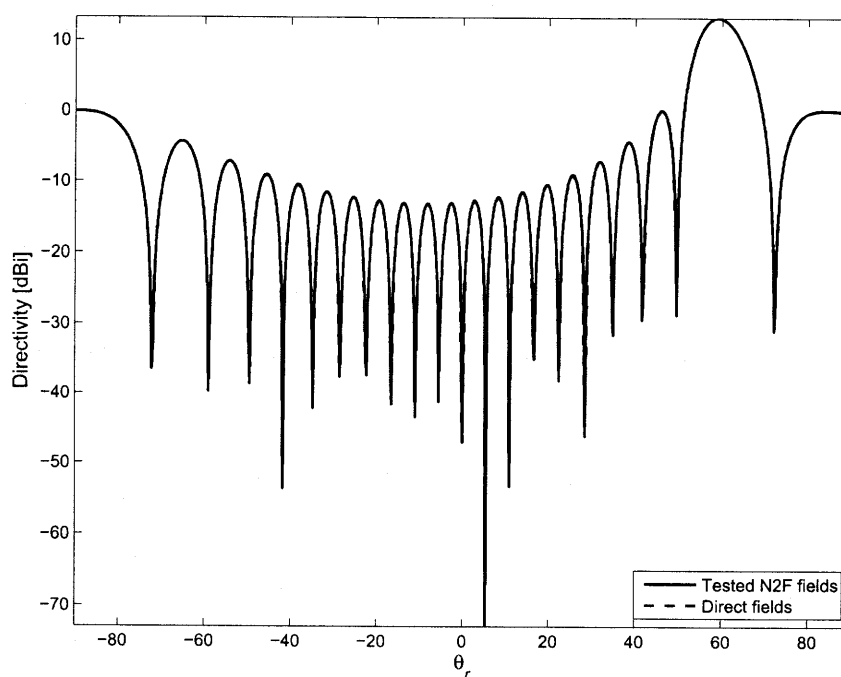

Figure 2.16: Pattern of the fields produced by a linear array of 21 point sources steered at 59.0625° and tested by a 17 vectors basis.

nearest the tested angle is from the end fire direction (see figure 2.17),

Conclusion 4   *The near fields produced by a beamsteering operation around the broad side direction require a higher order basis than the fields resulting from a beamsteering close to the end fire direction.*
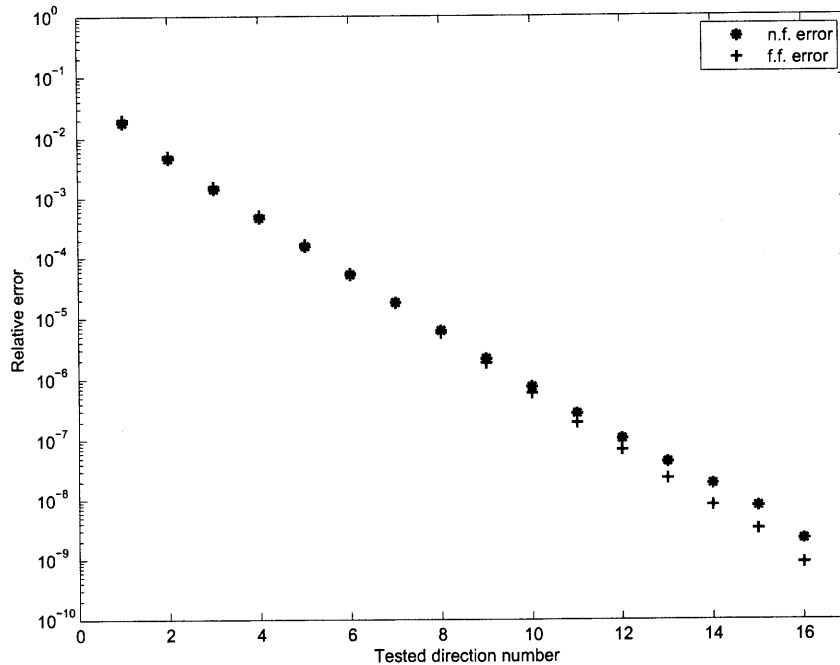

Figure 2.17: Near fields error and N2F error resulting from testing fields scanned in the middle of 17 exponentially selected scan angle spanning vectors for an array of 21 point sources.

Thus, we can choose the spanning vectors with higher density of the scanned angles around the broad side direction in order to achieve a constant error (in the average) along the whole scanning range (see figure 2.18).

Seeking for further improvements in the space selection, we have performed transformations on the near fields pictures before we collect them for $\mathbf{A}_\theta$. Exploiting the fact that the near fields are sampled on a sphere, we have rotated the fields around the y axis in order let the scan direction going back to the broadside direction.

The results with rotated fields are shown in figure 2.21. In figure 2.22 it is shown the results from a rotation in the direction in which the near fields are maximal. For
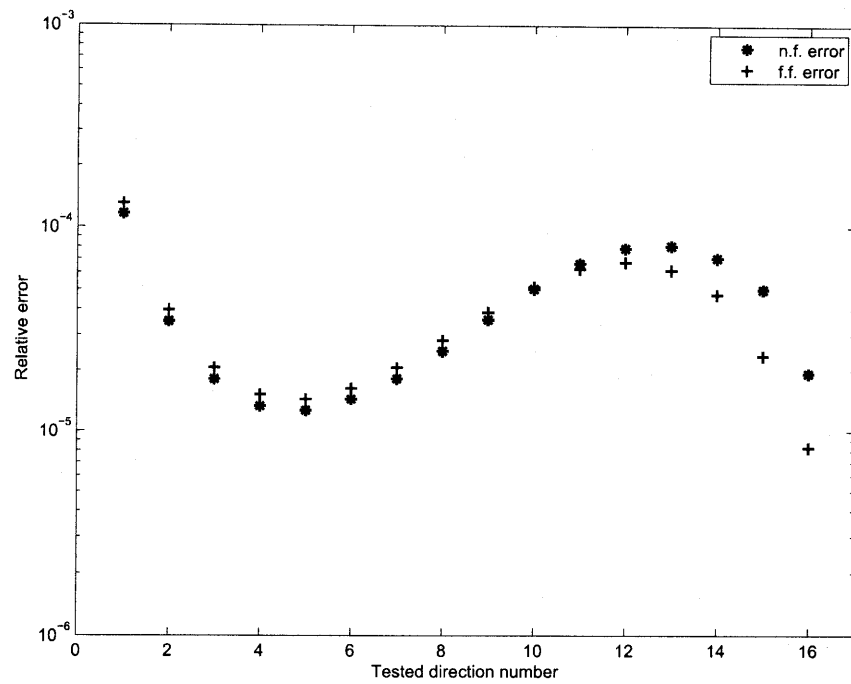
Figure 2.18: Possible correction applied to the selected angles for the spanning space. The gap between the angles is chosen to grow exponentially with base ≈ 1.1.
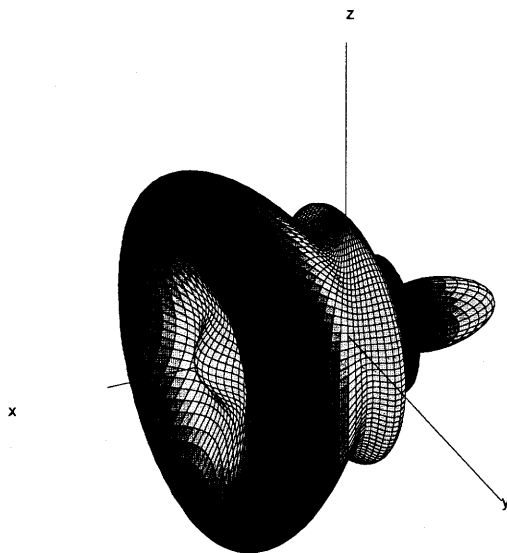

Figure 2.19: Near fields solid $|\psi|$ generated by a linear array of 9 point sources on the x direction steered at 30°.
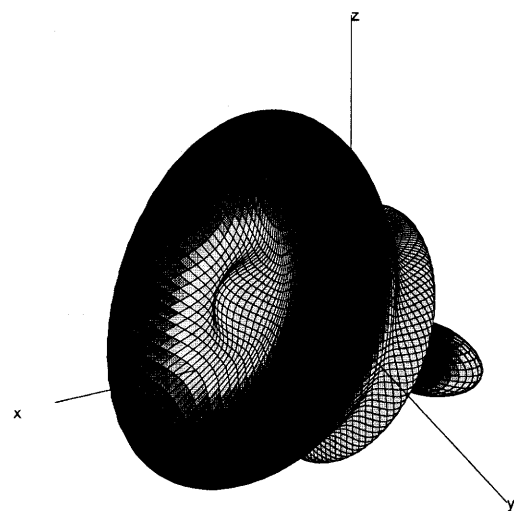
Figure 2.20: Near fields solid $|\psi|$ generated by a linear array of 9 point sources on the x direction steered at 30° rotated 30° backward.

this case, it is necessary, in order to have the computed maximal direction to be around the scan angle, to increase the size of the bounding sphere, such that the constructive interference becomes visibly in the scan angle direction (the behavior in the near fields is really different from the one in the far fields).
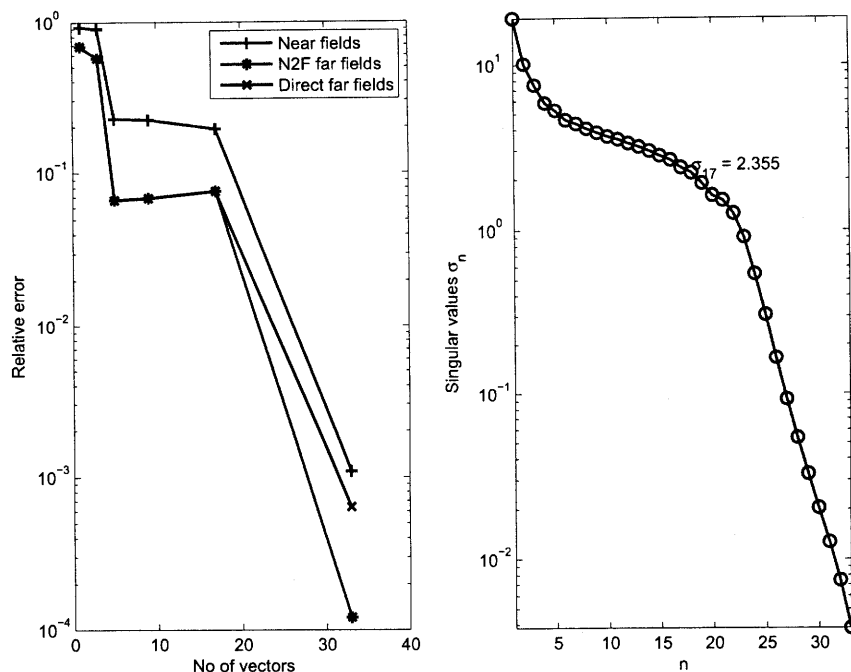
Figure 2.21: Error plot for the space spanned by fields rotated to broadside of the scan angle (linear array of 17 point sources).

As we can see, no improvement is achieved in this manner. Even the rank of the space is no more discernible, that is the spanning vectors are still very linearly independent one from each other, thus cannot constitute a basis for the space sought.

Interpolation / Approximation

2.2   Reduced basis for the near-to-far fields operator

↳ What you now present is more like a special kind of empirical interpolation

The far fields detectors presented in chapter 1 can be thought as operators [16] *to arrive at an* [17] acting on the near fields. In the scalar case, we can think of two (integral) opera- *affine parametrization* tors, the first acting on the scalar near fields and the second on their derivatives[4]. In

---

[4] In section 1.2.2.2, the derivatives where computed analytically in order to reduce the numerical errors that would be introduced by a numerical finite-differences-based derivation technique.
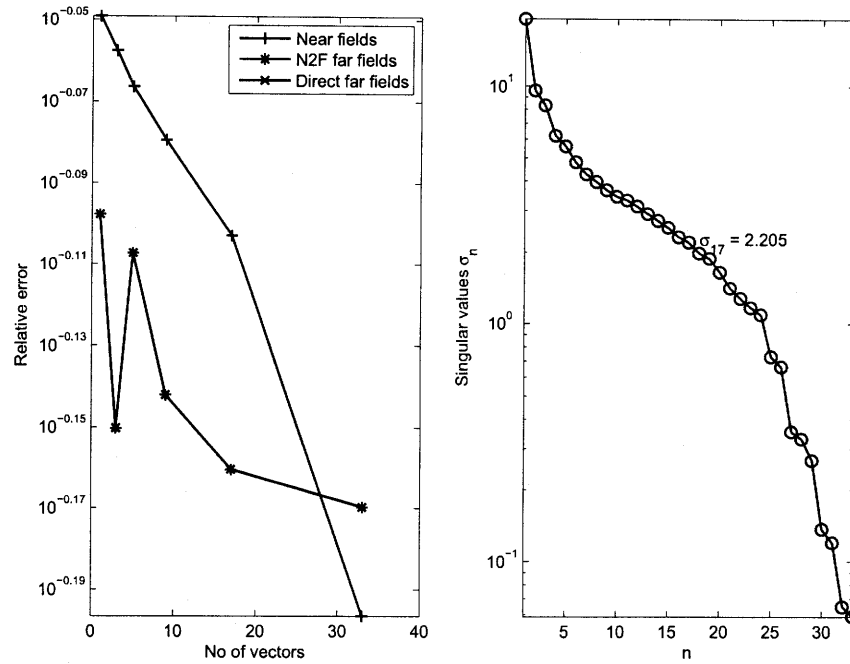
Figure 2.22: Error plot for the space spanned by fields rotated such that the maximum fields value goes to broadside (linear array of 17 point sources).

the vector case, six operators act on the components of the equivalent currents or of the near fields, modifying in the latter case the operators such that the normal unit vectors become part of them. In this section, we will treat only scalar fields to find *reduced basis approximations* for the N2F operator. However, the techniques that will be presented can be straightforwardly applied to vector fields. The far fields detector of equation (1.2.87) become, in such an interpretation,

$$
\tilde{\psi}_r^*(\hat{r};\omega) = \mathcal{L}_\psi\left(\tilde{\psi}(r';\omega);\hat{r}\right) + \mathcal{L}_{\psi'}\left(\frac{\partial \tilde{\psi}(r';\omega)}{\partial r'};\hat{r}\right)
$$

$$
= \mathcal{L}\left(\tilde{\psi}(r';\omega), \frac{\partial \tilde{\psi}(r';\omega)}{\partial r'};\hat{r}\right), \tag{2.2.5}
$$

where $\mathcal{L}$ is a linear[5] functional parameterized in the look angle direction $\hat{r} \in \mathcal{D} \subset \mathbb{R}^3$. In a first approach, we will restrain the look angle domain $\hat{r} \in \mathcal{D}_{\phi=0°} \subset \mathcal{D}$, seeking for a reduced basis approximation of $\mathcal{L}(\hat{r})$ along the pattern cut plane XZ ($\phi = 0°$).

However, we could have left the radial derivative to be part of the operator, leading into an integro-differential operator.

[5] The surfaces enclosing the sources can be more than one, an the far fields detector results in the sum of the application of the operator on each surface.

"reduced basis" usually refer to the ansatz space in MOR, i.e. the vectors you use for projection in the FE-context.

Here you are doing a special form of (empirical) interpolation.

The main idea of reduced basis is to construct basis functions with a large information content in order to reduce the number of basis coefficients needed to reach a certain level of accuracy in the outputs of interest. In a similar way as does the POD (section 2.1.1), we seek for the optimal basis to approximate the functional, in order to minimize the number of coefficients required.

2.2.1 Polynomial interpolation of the N2F operator

Given $\hat{r} \in \mathscr{D}_{\phi=0^\circ}$, we first introduce parameter samples $\mathbf{s}_P \equiv \{\hat{r}_1, \dots, \hat{r}_P\}$, that is the chosen look angles for pattern plotting, and associated *operator snapshots* $\mathbf{L}_P = \mathrm{colsp}\{\lambda_p \equiv \mathscr{L}(\hat{r}_p), 1 \le p \le P\}$. The column vector of far fields detectors ψ_r^* is given by the matrix vector product

$$\psi_r^* = \mathbf{L}_P^T \psi^{''} \tag{2.2.6}$$

where $\psi^{''} = [\psi; \psi']$ is the column vector of vertically concatenated near fields picture an̄d their radial derivatives, and $\mathbf{L}_P^T$ is the transposed matrix of $\mathbf{L}_P$. Numerically speaking, $\mathbf{L}_P$ is constructed modifying (1.2.94) in such a way that the summation becomes performed by the matrix multiplication (2.2.6). Thus we compute the values of the column vectors of $\mathbf{L}_P$ by the following

$$\mathbf{L}_p = \left[jk\left(\hat{r}_p \cdot \hat{r}'_m\right) e^{jk\hat{r}_p \cdot r'_m} \overset{\triangle}{d}S_m; -e^{jk\hat{r}_p \cdot r'_m} \overset{\triangle}{d}S_m \right],$$

↗ I would write ΔS_m since dS_m implies infinitesimal area element.

$$m = 1 \dots M, \ p = 1 \dots P, \tag{2.2.7}$$

and the concatenated vectors of the fields and their derivatives are

$$\psi^{''} = \left[\tilde{\psi}_m(r'_m; \omega); \frac{\partial \tilde{\psi}_m(r'_m; \omega)}{\partial r'_m} \right], \qquad m = 1 \dots M. \tag{2.2.8}$$

Looking first at the singular values computed from the factorization of $\mathbf{L}_P$ (figure 2.23) built for a sphere bounding a linear array of 5 point sources (radius of $3\lambda/2$), we see that the number of basis that may be required, at least for approximation through projection, is about 48. $\mathbf{L}_P$ This number is sensibly high, as we seek for a method that would be able to handle very large arrays (50 by 50 radiating elements and more). Thus, we proceed with a polynomial interpolation with the attempt to need less basis *vectors*
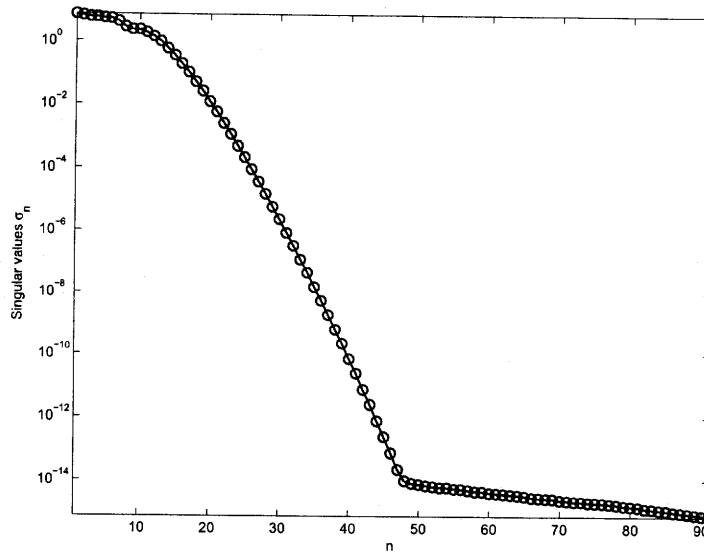
Figure 2.23: Singular values associated to the SVD of $\mathbf{L}_P$.

We introduce nested parameter samples $\mathbf{s}_Q \equiv \{\hat{r}_1, \ldots, \hat{r}_Q\}$ and associated approximation spaces $\mathbf{L}_Q = \text{colsp}\{\lambda_q \equiv \mathscr{L}(\hat{r}_q), 1 \leq q \leq Q\}$ for $Q = 1, \ldots, P$. Now, we evaluate $\mathscr{L}(\hat{r})$, assuming $\mathscr{L}$ "affine" in $\hat{r}$ by the following sum

$$\hat{\mathscr{L}}(\hat{r}) = \sum_{q=1}^{Q} \Theta_q(\hat{r}) \mathscr{L}(\hat{r}_q), \qquad (2.2.9)$$

have to be sample directions! otherwise they don't have to be 0.

$\Theta_m(\hat{r}_i) | \Theta_n(\hat{r}_j) = \partial$

where $\Theta_q(\hat{r})$ is a set of Q polynomial basis functions with compact supports $\mathscr{Q}_q \subset \mathscr{D}_{\phi=0°}$ such that $\underline{\Theta_m(\hat{r})\Theta_n(\hat{r}) = 0}$ if $m \neq n$, $\forall m, n = 1, \ldots, Q$. The affinity of $\mathscr{L}$ in $\hat{r}$, that is a polynomial dependence in the parameter, may actually, according to [21] *the affine* [22], allow us to define global smooth basis functions in order to reduce sensibly the *parameter dependence is actually* number of terms in the sum (2.2.9). This would lead into a very efficient approx- *a prerequisite* imation technique with *a posteriori* computable error bounds. Hence, we proceed *for the projection step in* checking the affinity of the operator in the look angle parameter space. *the MOR approach!*

Using *piece-wise linear* basis functions (1st order), we assume L_Q to be our interpolating nodes, and their influence is limited to the domain $\mathscr{Q}_q \in (\hat{r}_{q-1}, \hat{r}_{q+1})$ through *roof-top* shape functions as depicted in figure 2.24. For each node q, the roof-top function can be defined as, being $\hat{r} = \hat{r}(\theta)$ with $\theta = (\theta, \phi = 0°)$ the look

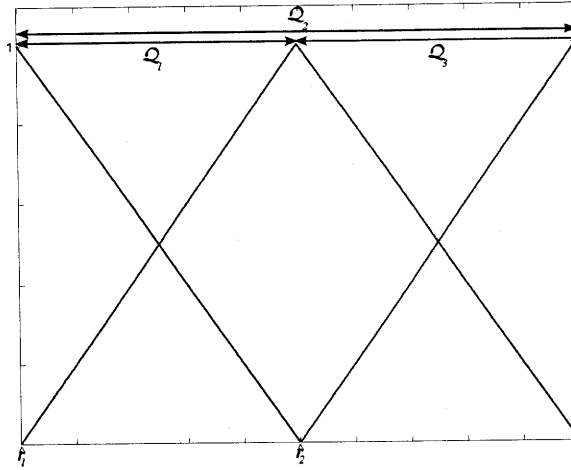Figure 2.24: *Roof-top* interpolating functions for the first 3 nodes. *[handwritten: plot θ on this axis, so one can establish a link to (2.2.10)]*

angle,

$$
\Theta_q(\hat{r}) \;=\;
\begin{cases}
\dfrac{\theta - \theta_{q-1}}{\theta_q - \theta_{q-1}}, & \hat{r}(\theta_{q-1}) \in \mathscr{D}_{\phi=0^\circ},\, \hat{r}(\theta) \in \mathscr{Q}_q, \\[2mm]
\dfrac{\theta_{q+1} - \theta}{\theta_{q+1} - \theta_q}, & \hat{r}(\theta_{q+1}) \in \mathscr{D}_{\phi=0^\circ},\, \hat{r}(\theta) \in \mathscr{Q}_q, \\[2mm]
0 & \text{otherwise.}
\end{cases}
\tag{2.2.10}
$$

Numerically speaking, we can build the interpolated operator selecting first in the range of P look angles Q of them equally spaced one from each other, building the matrix $\mathbf{L}_Q$ of Q columns. Then we build the matrix $\mathbf{O}_Q \in \mathbb{R}^{P \times Q}$, which column vectors are the weighting values given by the roof-top functions. Finally, the product of these matrices gives

$$
\hat{\mathbf{L}}_P = \mathbf{L}_Q \mathbf{O}_Q^T \approx \mathbf{L}_P.
\tag{2.2.11}
$$

Now, we can compute $\psi_r^*\big|_{1^{st}} = \hat{\mathbf{L}}_P^T \psi''$ and compare it to (2.2.6) in order to quantify the error introduced by the approximation of the operator. Figure 2.25 shows the errors in far fields detectors relatively to the full non approximated ones and to the direct fields ones. The $\mathbf{L}_P$ have been constructed with an average[6] 1000 look angles in the range of $\theta \in [-90^\circ, 90^\circ]$, for a sphere of radius $3\,\lambda/2$ bounding a linear array of 5 point sources equally spaced of $\lambda/2$, with $\lambda/10$ of sampling resolution ($\epsilon_{\text{Direct–N2F}[\lambda/10]} = 2.68\ 10^{-3}$). We notice that we need about 130 snapshots in order to introduce an

---

[6] As we require the basis look angles to be equally spaced, the full operator dimension result to be proportional to the number of look angles sampled.
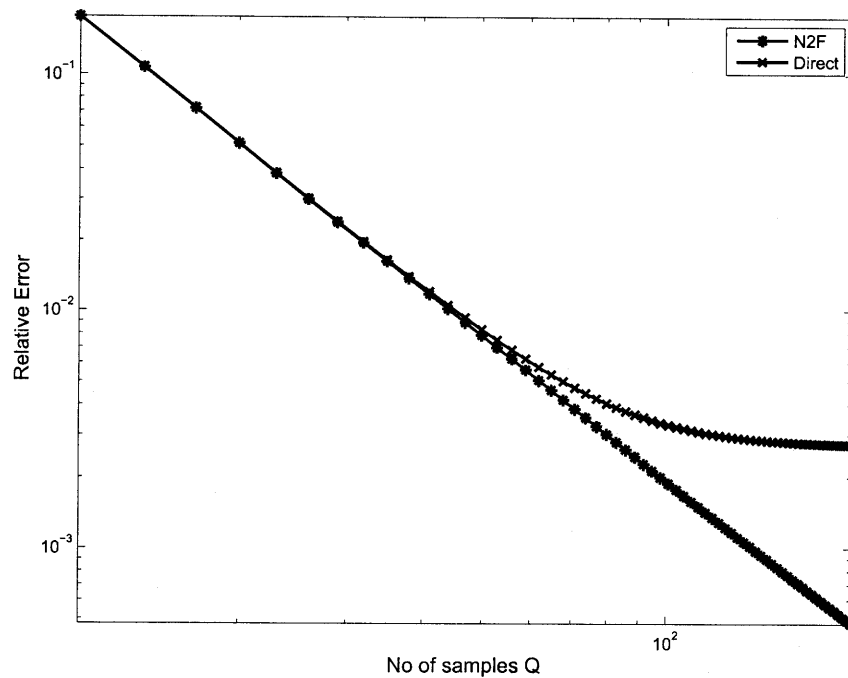
Figure 2.25: N2F error introduced with the *roof-top* interpolating functions.

error below 10^{-3}, such that the approximated operator appears to be non-influent on the far fields computation.

As the pattern is in general a rapidly variating (many orders of magnitude between constructive and destructive interferences) function of the look angle dependent on the number of radiators, many snapshots would be needed to approximate accurately a larger array than 5. We would need a higher slope in the errors as the number of snapshots increases, that is a quick convergence of the approximation into the exact operator. The finite element technique, also based on compact basis functions to approximate a global function, teaches us that we may expect faster convergence with higher order polynomials [23]. Thus, we proceed introducing the 2nd *order Lagrange polynomials* as depicted in figure 2.26. The 3 *quadratic* basis functions of figure 2.26 are such that the influence of a snapshot is slightly extended to the range in between two neighboring nodes, without affecting, as for the roof-top functions, the values at the neighboring sampled look angles. The basis of the first
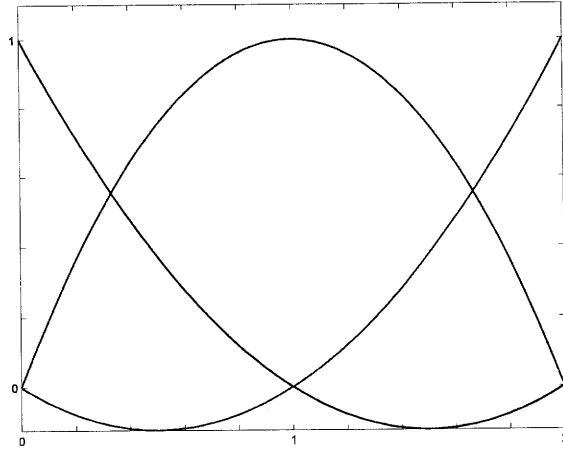
Figure 2.26: 2nd order interpolating polynomials for the first 3 nodes.

element (group of 3 nodes for quadratic basis functions) are expressed as,

$$\Theta_1(\hat{r}) = \begin{cases} (1-\xi)(2(1-\xi)-1), & 0 \text{ elsewhere} \\ 0 \end{cases} \quad \theta \in [\theta_1, \theta_3]$$

elsewhere

$$\Theta_2(\hat{r}) = 4(1-\xi)(\xi), \quad 0 \text{ elsewhere}$$

$$\Theta_3(\hat{r}) = \xi(2\xi-1), \quad 0 \text{ elsewhere} \quad \text{unusual notation}$$

$$\theta \in [\theta_1, \theta_3], \xi = \frac{\theta_3 - \theta}{\theta_3 - \theta_1}. \tag{2.2.12}$$

The sequence is then repeated for $q = 4$, 5 and 6 with the new range of $\theta \in [\theta_4, \theta_6]$ and $\frac{\theta_6 - \theta}{\theta_6 - \theta_4}$, and so on for $q > 6$. We proceed building the approximated operator, and compute the far fields detectors $\psi_r^*\big|_{2^{nd}}$. The sampled look angles define the full operator look angles, in order to preserve equal spacing between the interpolation angles. The error convergence plot is shown in figure 2.27. We can see that we get below 10^{-3} with only about 50 snapshots, compared to the ≈ 130 required by the roof-top approximation. For 200 snapshots, we notice that we achieve an error one order of magnitude lower.

In order to see if we can achieve even faster convergence, 3rd *order Lagrange poly-nomials* have been also tested. Figure 2.28 shows the *cubic* basis of an element, and they are analytically expressed as

$$\Theta_1(\hat{r}) = \frac{1}{2}(1-\xi)(2-3\xi)(1-3\xi), \quad 0 \text{ elsewhere}$$

$$\Theta_2(\hat{r}) = \frac{9}{2}\xi(1-\xi)(2-3\xi), \quad 0 \text{ elsewhere}$$
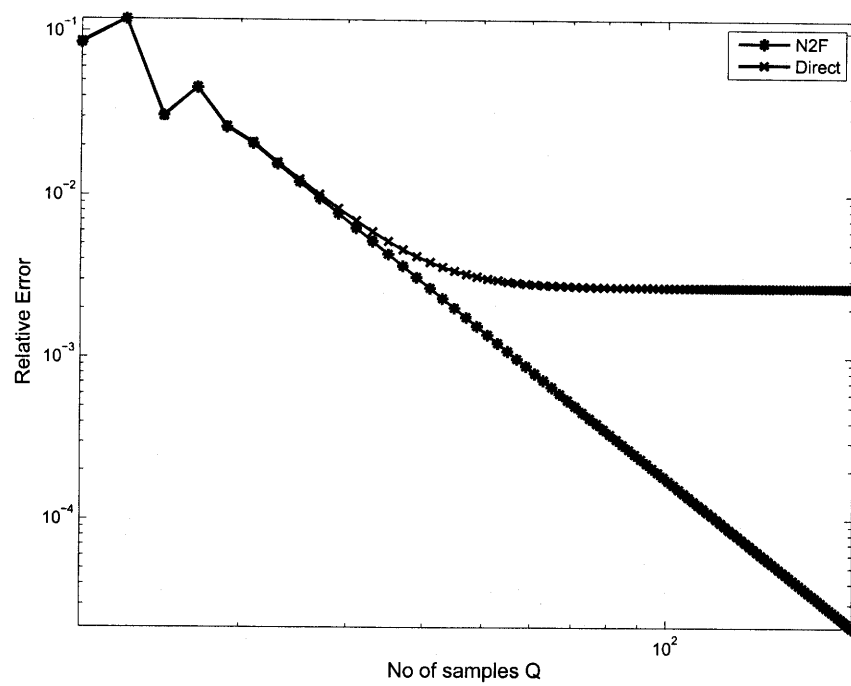
Figure 2.27: N2F error introduced with the 2nd order interpolating functions.

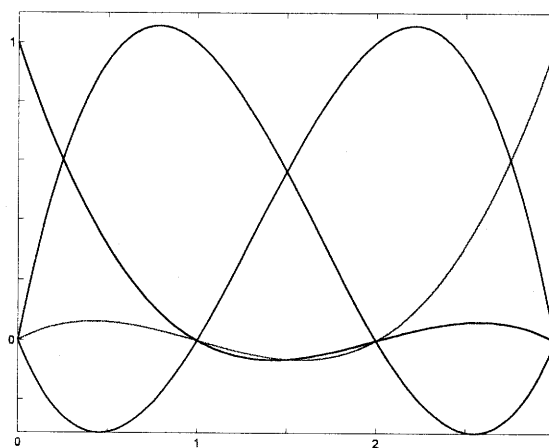Figure 2.28: 3rd order interpolating polynomials for the first 4 nodes.

$$\Theta_3(\hat{r}) = \frac{9}{2}\xi(1-\xi)(3\xi-1), \ 0 \text{ elsewhere}$$

$$\Theta_4(\hat{r}) = \frac{1}{2}\xi(2-3\xi)(1-3\xi), \ 0 \text{ elsewhere}$$

$$\theta \in [\theta_1, \theta_4], \xi = \frac{\theta_4 - \theta}{\theta_4 - \theta_1}. \tag{2.2.13}$$

The results of the approximation, made similarly to the two previous cases, are shown in figure 2.29. The convergence slope is higher. Only about 35 snapshots are
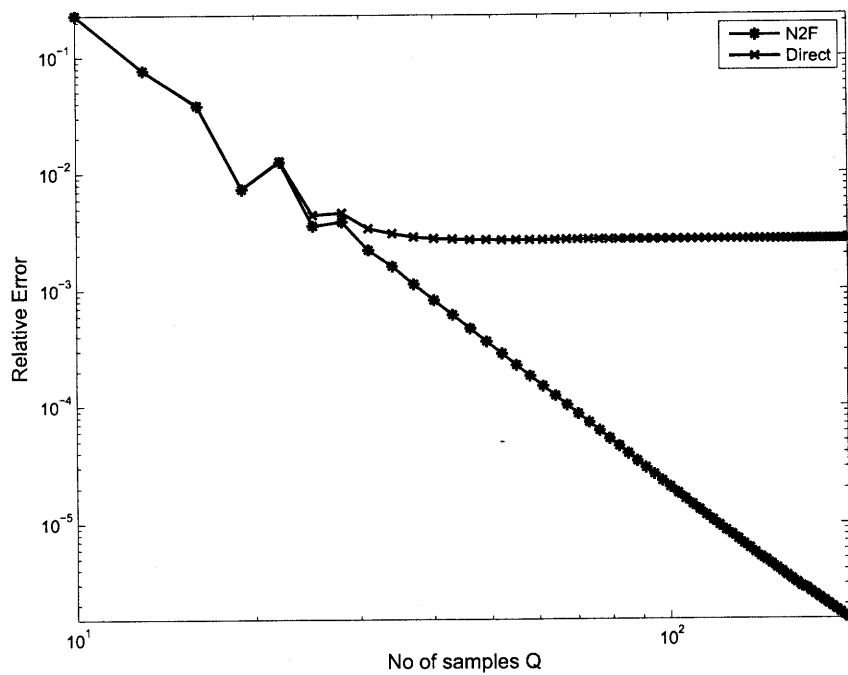

Figure 2.29: N2F error introduced with the 3$^{\text{rd}}$ order interpolating functions.

needed to achieve an error below 10^{-3}. With 200 snapshots, we have achieved an error two orders of magnitude below the one we had with roof-top functions. Extrapolating the convergence behaviors for each of the error plots of figures 2.25, 2.27 and 2.29, we may conclude that

Conclusion 5 *Standard FEM 1$^{\text{st}}$, 2$^{\text{nd}}$, 3$^{\text{rd}}$ order polynomials allow the approximation of the N2F operator in the look angle parameter, and the convergence is in the order of, respectively,* 40 $^{\text{dB}}$/dec, 60 $^{\text{dB}}$/dec *and* 80 $^{\text{dB}}$/dec.

The number of snapshots required by a 3rd order basis for the interpolation of an operator acting on the fields produced by an array of 5 point sources is about 35. Before we proceed in the trigonometric polynomials approximation, we check how the error increases as we further increase the radius of the sphere letting double the number of the point sources. The results are shown in figure 2.30. The number of
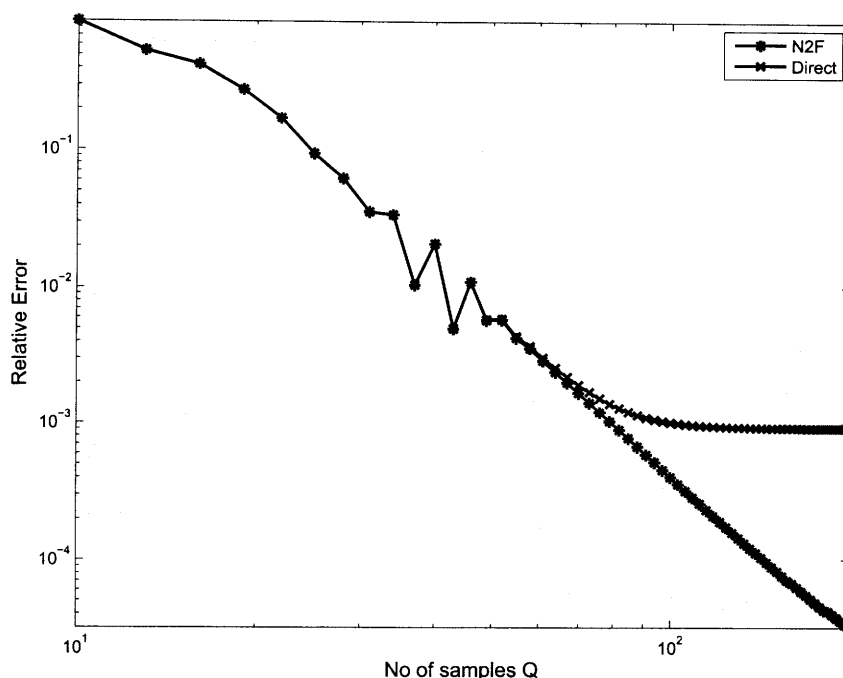

Figure 2.30: N2F error introduced with the 3rd order interpolating functions for an operator build on a sphere enclosing 10 point sources (radius of $5\lambda/4$ and sampling resolution of $\lambda/10$).

snapshots required to preserve the accuracy is about 100, a value that is too high to consider approximating very large arrays.

2.2.2 Trigonometric polynomials interpolation of the N2F Operator

Using a similar formalism as introduced the previous sections 2.1.1 and 2.2.1, we extend the notion of polynomial basis approximation to the case of *complex trigono-*

metric polynomials of the form

$$z(r(\theta)) = \frac{a_0}{2} + \sum_{m=1}^{M} [a_m \cos(m\theta) + b_m \sin(m\theta)],$$
$$a_m, b_m \in \mathbb{C} \text{ for } 0 \le m \le M, \quad \theta \in \mathbb{R}, \tag{2.2.14}$$

where we have assumed $\hat{r} \in \mathcal{D}_{\phi=0°}$. Using *Euler's formulas* with $e^{im\theta} = \cos(m\theta) + i\sin(m\theta)$, i being the imaginary unit, (2.2.14) can be rewritten as [26]

$$z(r(\theta)) = \sum_{m=-M}^{M} c_m e^{im\theta}, \quad c_m \in \mathbb{C} \text{ for } -M \le m \le M, \quad \theta \in \mathbb{R}. \tag{2.2.15}$$

Now, the complex Fourier series approximation of a function $f(x)$, periodic[7] of period L such that $f(x) = f(x + L)$, is given by

$$f(x) = \sum_{r=-\infty}^{\infty} c_r e^{i\frac{2\pi}{L} rx}, \tag{2.2.16}$$

and the complex coefficients c_r are

$$c_r = \frac{1}{L} \int_{x_0}^{x_0+L} f(x) e^{-i\frac{2\pi}{L} rx} dx, \quad x_0 \in \mathbb{R}. \tag{2.2.17}$$

$\phi_r(x) = e^{i\frac{2\pi}{L} rx}, \forall r \in \mathbb{Z}$ is a set of orthogonal functions or *eigenfunctions* in the sense of (2.1.2). Now, considering $x = \theta$, and as the domain of θ is, for the spherical geometry of the radiation problem, periodic of $360° = 2\pi$, we set $L = 2\pi$. For $M \to \infty$, we have that the finite sum approximation (2.2.15) becomes the Fourier approximation stated in (2.2.16) with the variables changes made, thus we can use the highly developed mathematical and numerical foundations on the Fourier series to compute an approximant of $z(r(\theta))$. As we are interested in an approximation of the operator $\mathscr{L}(\hat{r}(\theta))$, we shall assume $\mathscr{L}(\hat{r}(\theta))$ affine in trigonometric polynomials and evaluate $z(r(\theta)) = \hat{\mathscr{L}}(\hat{r}(\theta))$ the interpolated functional.

The coefficients c_r of (2.2.17) may be computed by a *Discrete Fourier Transform* (DFT), i.e. by a numerical integration of (2.2.17), the interval $[x_0, x_0 + L)$ being subdivided into N equal intervals, and the N samples of $f(x)$ taken in a domain of zero

---

[7] The first of the four Dirichlet conditions [25].

measure[8]. This step can be efficiently performed through the *Fast Fourier Transform* (FFT) algorithm [26]. The FFT computes the Fourier coefficients $\mathbf{X}(m)$ of a sequence $\mathbf{x}(n)$, assuming $\mathbf{x}(n)$ to be repeated infinitely and periodically along the sampling space and performing the following summations:

$$\mathbf{X}(m) = \sum_{n=1}^{N} \mathbf{x}(n) e^{-i\frac{2\pi}{N}(n-1)(m-1)}, \qquad m = 1,...,N. \qquad (2.2.18)$$

The assumption of the periodicity of $\mathbf{x}(n)$, in our case, is valid as long as our operator is built by scanning θ all over 360°. The sampled sequence can be recovered from the coefficients by the *Inverse FFT* (iFFT)

$$\mathbf{x}(n) = \frac{1}{N} \sum_{m=1}^{N} \mathbf{X}(m) e^{\frac{i2\pi}{N}(n-1)(m-1)}, \qquad n = 1,...,N. \qquad (2.2.19)$$

Now, assuming the N2F operator to be $\mathscr{L}(\hat{r}(\theta)) = \mathscr{L}(\theta)$, and $\mathbf{x}(n)$ the sequence obtained by the snapshots

$$\mathbf{x}(n) = \mathscr{L}(\theta_n) = \mathscr{L}\left(\frac{2\pi}{N}(n-1)\right), \qquad n = 1,...,N,$$

we can express $\mathscr{L}(\theta_n)$ by means of the Fourier coefficients $\mathbf{X}(m), m = 1,...,N$ with their respective eigenfunction $e^{j\theta_n(m-1)}$, $m-1$ being the associated *eigenfrequency*. For the FFT algorithm, the spectrum of the eigenfrequencies is such that the first $m = 2,...,\lfloor N/2 \rfloor + 1$ correspond to positive frequencies $(m-1)$ and the next $m = \lfloor N/2 \rfloor + 2,...,N$ correspond to the negative frequencies $N + 1 - m$. The continuous spectral component corresponds to $m = 1$. *Aliasing* [26] in the original sequence reconstruction may occur, and for the *sampling theorem* we need to sample the operator such that $N \geq 2q$ where q is the highest eigenfrequency associated to non null (or numerically significant) coefficient values $\mathbf{X}(q)$.

The discretized eigenfunctions of the FFT can be extrapolated to the continuous case considering $\lim_{N\to\infty} \theta_n = \theta \in [0, 2\pi)$, and we obtain $\mathscr{L}(\theta_n) \to \mathscr{L}(\theta)$. Thus, we can express the N2F operator as, if the number of sampling points N is odd,

$$\mathscr{L}(\theta) = \frac{1}{N}\left(\mathbf{X}(1) + \mathbf{X}(2)e^{i\theta} + \mathbf{X}(3)e^{i2\theta} + \ldots\right.$$

---

[8] Zero measure for the sampled function is achieved by a Dirac comb, simplifying the integral by the sifting property of the Dirac delta.

$$\ldots + \mathbf{X}(\lfloor N/2 \rfloor + 1)e^{i \lfloor N/2 \rfloor \theta} + \mathbf{X}(\lfloor N/2 \rfloor + 2)e^{-i(\lfloor N/2 \rfloor + 1)\theta} + \ldots$$
$$\ldots + \mathbf{X}(N-2)e^{-i3\theta} + \mathbf{X}(N-1)e^{-i2\theta} + \mathbf{X}(N)e^{-i\theta}\Big), \qquad (2.2.20)$$

or as, if the number of sampling points N is even,

$$\mathscr{L}(\theta) = \frac{1}{N}\Big(\mathbf{X}(1) + \mathbf{X}(2)e^{i\theta} + \mathbf{X}(3)e^{i2\theta} + \ldots$$
$$\ldots + \mathbf{X}(N/2)\cos(\frac{N}{2}\theta) + \ldots$$
$$\ldots + \mathbf{X}(N-2)e^{-i3\theta} + \mathbf{X}(N-1)e^{-i2\theta} + \mathbf{X}(N)e^{-i\theta}\Big). \qquad (2.2.21)$$

Now, performing the trigonometric polynomials approximation or *DFT-truncation* by limiting the number of coefficients to $R \in \mathbb{N}$ around the continuous spectral component, that is the odd number of coefficients $2R+1$, the full interpolations of (2.2.20) and (2.2.21) reduce to

$$\mathscr{L}(\theta) = \frac{1}{N}\Big(\mathbf{X}(1) + \mathbf{X}(2)e^{i\theta} + \mathbf{X}(3)e^{i2\theta} + \ldots$$
$$\ldots + \mathbf{X}(R+1)e^{iR\theta} + \mathbf{X}(N-R+1)e^{-iR\theta} + \ldots$$
$$\ldots + \mathbf{X}(N-2)e^{-i3\theta} + \mathbf{X}(N-1)e^{-i2\theta} + \mathbf{X}(N)e^{-i\theta}\Big). \qquad (2.2.22)$$

It is now possible to apply the DFT-truncation to a selected N2F problem. We chose, in order to compare the results to those obtained by the FEM basis based approximation, a linear array of 5 point sources within a sphere of radius of $3\lambda/2$ and sampling resolution of $\lambda/10$. The number of look angles in order to build the operator matrix of snapshots $\mathbf{x}_N$ is of $N = 300$ in the range [0°, 360°). Figure 2.31 shows the spectral coefficients $\mathbf{X}$ computed by an FFT on the rows (look angle dependent) of $\mathbf{x}_N$.

The spectrum is confined within ±32 coefficients around the continuous spectral component, range over which they remain in the numerical null level (≈ -250dB).

The computation of the error induced in the pattern by the DFT-truncation is shown in figure 2.32. As we can see, the error drops rapidly from 11 coefficients till we have null error with only 65 coefficients. For our purposes, we can retain 21 coefficients in order to achieve the required accuracy in the pattern. This value is lower
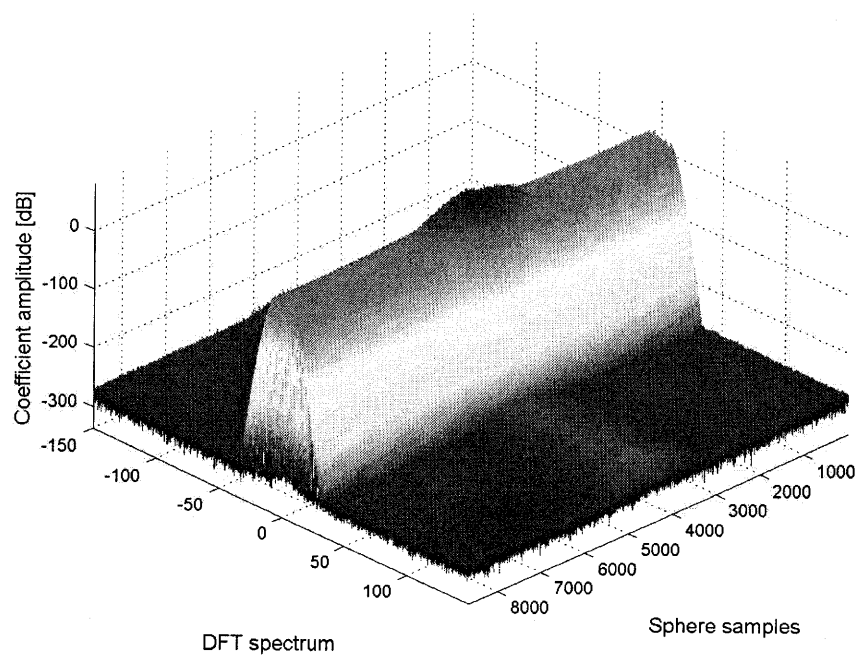
Figure 2.31: Spectral coeffients of an operator build for a sphere enclosing 5 point sources (radius of $3\lambda/2$ and sampling resolution of $\lambda/10$).
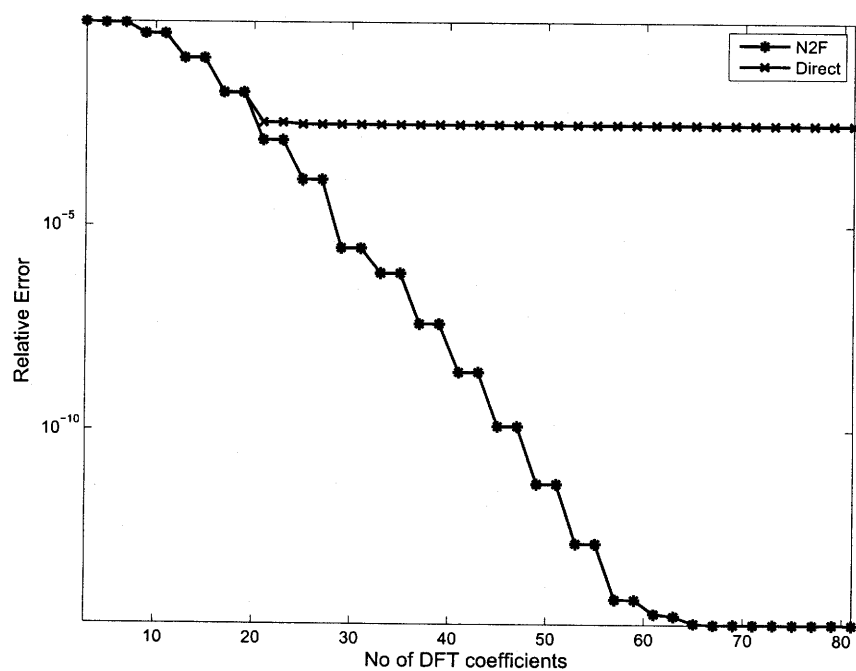

Figure 2.32: Error in the far fields induced by DFT-truncation.

than that of the 3rd order basis polynomial interpolation, where we needed about 35 snapshots.

We may now check the performances of the DFT-truncation as the array grows in the number of point sources. We choose 10 point sources in order to compare the performances to the 3rd order basis approximation results of figure 2.30. The results are depicted in figure 2.33. Only 41 coefficients are needed to accurately perform
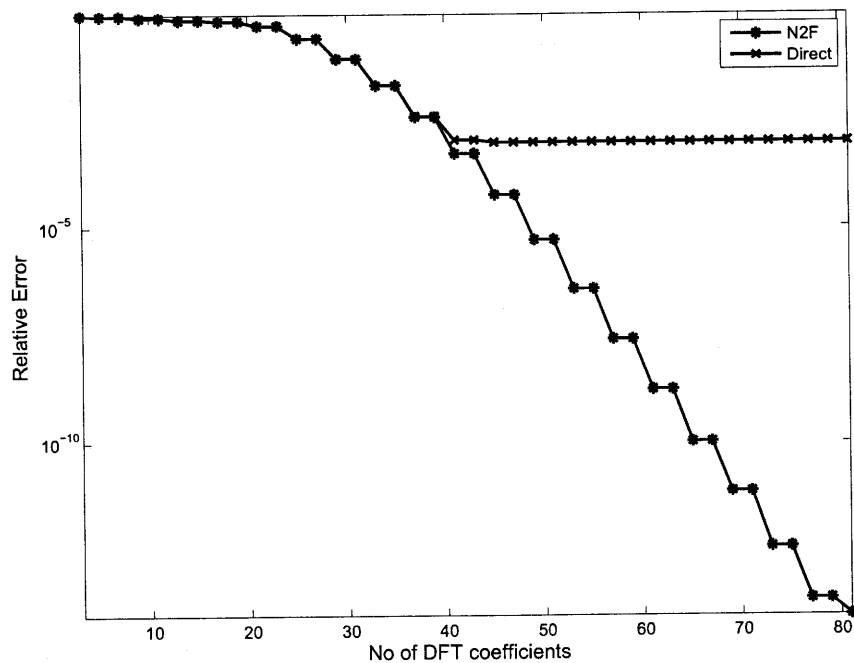

Figure 2.33: Error in the far fields induced by DFT-truncation for the linear array of 10 point sources.

the pattern computations, which is definitely lower than the 100 required for the 3rd order basis approximation.

As we have seen along the process of choosing the appropriate approximation space for the N2F operator, we can conclude that

Conclusion 6 *The relevant Fourier coefficients have resulted to be a few ones, allowing us to represent the N2F operator as a finite sum of trigonometric polynomials. The performances of the DFT-truncation are better than those achieved by the polynomial interpolation. To compute the values of the coefficients, it is required to oversample the*

operator in order to avoid aliasing, and this step is dependent on the size of the array enclosed which determines the highest eigenfrequency. The DFT-truncation technique is very promising for large arrays approximations.