Pré-Publicações do Departamento de Matemática
Universidade de Coimbra
Preprint Number 13–35

MEASURES OF LOCALIZATION AND QUANTITATIVE NYQUIST DENSITIES

LUÍS DANIEL ABREU AND JOÃO MORAIS PEREIRA

Abstract: We obtain quantitative versions of the Nyquist density, by estimating, in terms of ϵ, the increase in the degrees of freedom resulting upon allowing the functions to contain a certain prescribed amount of energy ϵ outside a region delimited by a set in time and a set in frequency. At the technical level, we study a *pseudospectra* version of the classical spectral problem of Landau-Slepian-Pollak. Analogue results are obtained for Gabor localization operators in a compact region of the time-frequency plane.

1. Introduction

1.1. The Nyquist rate. Let $D_{[-T,T]}$ and $B_{[-\Omega,\Omega]}$ denote the operators which cut the time content outside $[-T,T]$ and the frequency content outside $[-\Omega,\Omega]$, respectively. In a nowadays classical paper [15], whose purpose was to "*examine the true in the engineering intuition that there are approximately* $2T\Omega/\pi$ *independent signals of bandwidth* Ω *concentrated on an interval of length* T", Landau and Pollak have considered the eigenvalue problem

$$D_{[-T,T]}B_{[-\Omega,\Omega]}D_{[-T,T]}f = \lambda f. \tag{1.1}$$

The operator involved in this problem can be written explicitly as

$$(D_{[-T,T]}B_{[-\Omega,\Omega]}D_{[-T,T]}f)(x) = \begin{cases} \int_{-T}^{T} \frac{\sin\Omega(x-t)}{\pi(x-t)} f(t) & if \quad |x| < T \\ 0 & if \quad |x| > T \end{cases}$$

Received August 9, 2013.

L.D. Abreu was supported by CMUC through European program COMPETE/FEDER under the project PEst-C/MAT/UI0324/2011 and by FCT project PTDC/MAT/114394/2009 and START-project FLAME ('Frames and Linear Operators for Acoustical Modeling and Parameter Estimation'; Y 551-N13).

J. M. Pereira was supported by a FCG grant within the Gulbenkian program "Novos talentos em Matematica",

The cornerstone of the results in [15] is the following asymptotic estimate for the number of eigenvalues λ_n of (1.1) which are close to one:

$$\#\{n : \lambda_n > 1 - \delta\} = \frac{2T\Omega}{\pi} + \log T, \tag{1.2}$$

as $T \to \infty$. Since the eigenvalues of the equation (1.1) are the same as those of the operator $B_{[-\Omega,\Omega]}D_{[-T,T]}$, whose eigenfunctions satisfy

$$\int_{[-T,T]} |f|^2 = \lambda \|f\|^2,$$

the estimate (1.2) provides us with the number of orthogonal eigenfunctions of (1.1) such that

$$\int_{[-T,T]} |f|^2 \sim \|f\|^2,$$

asymptotically when $T \to \infty$. Within mathematical signal analysis (see, for instance the discussion in [5, pag. 23] and the recent book [12]), (1.2) is viewed as a mathematical formulation of the Nyquist rate, the fact that a time- and bandlimited region $[-T, T] \times [-\Omega, \Omega]$ corresponds to $2T\Omega/\pi$ "degrees of freedom". In other words, there exist, up to a small error, $2T\Omega/\pi$ independent functions that are essentially timelimited to $[-T, T]$ and bandlimited to $[-\Omega, \Omega]$.

1.2. Quantitative versus qualitative. The main goal of this note is to formulate the concentration problem in a more quantitative way and to prove the resulting quantitative versions of (1.2).

Let us be more precise in order to define what we mean by "quantitative"; ideally, one would like to count the number of orthogonal functions in $L^2(\mathbb{R})$, which are time and band-limited to a bounded region like $[-T, T] \times [-\Omega, \Omega]$. Unfortunately, such functions do not exist (because band-limited functions are analytic). All we can do is to count the number of orthogonal functions in $L^2(\mathbb{R})$ which are approximately time and band-limited to a bounded region like $[-T, T] \times [-\Omega, \Omega]$. An optimal solution to this problem is given by the number of eigenfunctions of (1.1) whose *eigenvalues are very close to one*. The number of eigenfunctions with such a *quality* is given by (1.2). Thus, the eigenvalue approach can be seen as leading towards a *qualitative* solution of the concentration problem. In contrast, a *quantitative* solution would tell us the number of orthogonal functions in $L^2(\mathbb{R})$ with a prescribed *quantity ϵ of time-frequency content outside the bounded region* $[-T, T] \times [-\Omega, \Omega]$. In

order to do so, we propose to count the number of orthogonal functions in $L^2(\mathbb{R})$, ϵ-*localized* in the sense that

$$\left\| D_{[-T,T]} B_{[-\Omega,\Omega]} D_{[-T,T]} f - f \right\| \leq \epsilon. \tag{1.3}$$

From our main result it follows that (1.2) has the following analogue in this setting: if $\eta(\epsilon, [-T,T], [-\Omega,\Omega])$ stands for the maximum number of orthogonal functions of $L^2(\mathbb{R})$ satisfying (1.3), then, as $T \to \infty$,

$$\frac{2T\Omega}{\pi}(1+\epsilon^2)+\log T \leq \eta(\epsilon, [-T,T], [-\Omega,\Omega]) \leq \frac{2T\Omega}{\pi}(1-4\epsilon^2)^{-1}+\log T. \tag{1.4}$$

1.3. Localization operators. Our understanding of the concentration problem is based on the study of operators which localize signals in bounded regions of the time-frequency plane. Such operators are known in a broad sense as time-frequency localization operators; their eigenfunctions are orthogonal sequences of functions with optimal concentration properties. The quantitative formulation of the concentration problem can be seen in terms of localization operators as follows: rather than looking for the optimal concentrated functions in a given region of the time-frequency plane, we will allow the functions to contain a certain prescribed amount of energy outside the given region, and estimate the resulting increase in the degrees of freedom. Given an operator L, instead of counting the eigenfunctions of

$$Lf = \lambda f$$

associated with eigenvalues λ close to one, we will count orthogonal functions ϵ-*localized* with respect to L in the sense that

$$\|Lf - f\| \leq \epsilon. \tag{1.5}$$

1.4. Pseudospectra and $\epsilon-$localization. The result of Landau and Pollak has later been improved by Landau to several dimensions and more general sets than intervals in [13] and [14]. Also in [14], Landau introduced the concept of ϵ-approximated eigenvalues and eigenfunctions. This concept is a forerunner of what is nowadays known as the *pseudospectra* in the numerical analysis of non-normal matrices [18]. Recent developments in spectral approximation theory involve the concept of n-pseudospectrum, which has been introduced in [11] with the purpose of approximating the spectrum of bounded linear operators on an infinite dimensional, separable Hilbert space, and then used in the proof of the computability of the spectrum of a linear

operator on a separable Hilbert space [10]. We will recall Landau's original definition, which was the following:

Definition 1. *λ is an ϵ-approximate eigenvalue of L if there exists f with $\|f\| = 1$, such that $\|Lf - \lambda f\| \leq \epsilon$. We call f an ϵ-approximate eigenfunction corresponding to λ.*

Thus, our quantitative measure (1.5) for the time-frequency localization of f is equivalent to f being a ϵ-approximate eigenfunction corresponding to 1.

Example 1. *Consider functions satisfying*

$$\|D_{rT}B_\Omega D_{rT}f - f\| \leq \epsilon, \tag{1.6}$$

Then, ϵ measures how well the function is concentrated in the frequency band set Ω and in the time set rT. Suppose that φ is an eigenfunction of $D_{rT}B_\Omega D_{rT}$ with eigenvalue λ. Then

$$\|D_{rT}B_\Omega D_{rT}f - f\| = 1 - \lambda.$$

Thus, every eigenfunction of $D_{rT}B_\Omega D_{rT}$ is a $(1-\lambda)$-pseudoeigenfunction with pseudoeigenvalue 1.

The relevant fact is that there are more pseudoeigenfunctions with pseudoeigenvalue 1. A large class of functions satisfying (1.2) arises from the set of almost bandlimited functions in the sense of Donoho-Stark's concept of ϵ-concentration.

Example 2. *According to* [7]*, f is ϵ_T-concentrated in T if*

$$\|D_T f - f\| \leq \epsilon_T$$

and Tf is ϵ_Ω-concentrated in Ω if

$$\|B_\Omega f - f\| \leq \epsilon_\Omega. \tag{1.7}$$

An application of the triangle inequality shows that if f is ϵ_T-concentrated in T and Tf is ϵ_Ω-concentrated in Ω then

$$\|B_\Omega D_T f - f\| \leq \epsilon_T + \epsilon_\Omega.$$

and another application of the triangle inequality gives

$$\|D_T B_\Omega D_T f - f\| \leq 2\epsilon_T + \epsilon_\Omega. \tag{1.8}$$

Thus, if f is ϵ_T-concentrated in T and Tf is ϵ-concentrated in Ω, then f is an $(2\epsilon_T + \epsilon_\Omega)$-localized with respect to $D_T B_\Omega D_T$.

[11] A. C. Hansen, *On the approximation of spectra of linear operators on Hilbert spaces*, J. Funct. Anal. 254, (2008), 2092–2126.
[12] J. A. Hogan, J. D. Lakey, *Duration and Bandwith Limiting. Prolate Functions, Sampling, and Applications*, Applied and Numerical Harmonic Analysis, Birkhäuser/Springer, New York, 2012, xvii+258pp.
[13] H. J. Landau, *Necessary density conditions for sampling and interpolation of certain entire functions*, Acta Math., 117 (1967), 37-52.
[14] H. J. Landau, *On Szegö's eigenvalue distribution theorem and non-Hermitian kernels*, J. d'Analyse Math. 28 (1975), 335-357.
[15] H. J. Landau, H. O. Pollak, *Prolate spheroidal wave functions, Fourier analysis and uncertainty-III: The dimension of the space of essentially time- and band-limited signals*, Bell Syst. Tech. J., 41 (1962), 1295-1336.
[16] J. Ramanathan, P. Topiwala, *Time-frequency localization and the spectrogram*, Appl. Comput. Harm. Anal., 1 (1994), 209-215.
[17] K. Seip, *Reproducing formulas and double orthogonality in Bargmann and Bergman spaces*, SIAM J. Math. Anal. 22, 3 (1991), 856-876.
[18] L. N. Trefethen, M. Embree, *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators,* Princeton University Press, 2005.

Luís Daniel Abreu
Austrian Academy of Sciences, Acoustics Research Institute, Wohllebengasse 12-14, Vienna A-1040, Austria. On leave from CMUC, Department of Mathematics, University of Coimbra, Portugal

João Morais Pereira
Department of Mathematics, University of Coimbra, Portugal

Moreover, given $0 < \gamma < 1$, the number $M_r(\gamma)$ of eigenvalues which are not smaller than γ, satisfies, as $r \to \infty$,

$$M_r(\gamma) = r^d \, |S| + o(r^d).$$

Now that we have described the Gabor set-up in a close analogy to the band- time- limiting case, we obtain an analogue of Theorem 1 by performing minor adaptations in the proof.

Theorem 2. *Let $\eta(\epsilon, rS)$ stand for the maximum number of orthogonal functions $F \in L^2(\mathbb{R}^{2d})$ such that*

$$\|D_{rS}\mathcal{P}_g D_{rS} F - F\| \leq \epsilon. \tag{3.3}$$

Then, as $r \to \infty$, the following inequalities hold:

$$|S| \, (1 + \epsilon^2) \leq \lim_{r\to\infty} \frac{\eta(\epsilon, rS)}{r^{2d}} \leq \frac{|S|}{1 - 4\epsilon^2}.$$

Proof: The proof mimics the proof of Theorem 1, replacing $D_{rT}B_\Omega D_{rT}$ by $D_{rS}\mathcal{P}_g D_{rS}$, $B_\Omega D_{rT}$ by $\mathcal{P}_g D_{rS}$ and Proposition 1 by Proposition 2. ∎

Acknowledgement. The authors thank José Luís Romero for his constructive cricticism of earlier versions of the manuscript, leading to a better formulation of the results.

References

[1] L. D. Abreu, A. S. Bandeira, *Landau's necessary density conditions for the Hankel transform*, J. Funct. Anal. 162 (2012), 1845-1866.
[2] L. D. Abreu, M. Dörfler, *An inverse problem for localization operators*, Inverse Problems, 28 (2012), 16 pp, 115001.
[3] E. Cordero and K. Gröchenig. *Time-frequency analysis of localization operators*, J. Funct. Anal. 205 (2003), 107–131.
[4] I. Daubechies. *Time-frequency localization operators: a geometric phase space approach*, 34 (1988), IEEE Trans. Inform. Theor. 605–612.
[5] I. Daubechies, *"Ten Lectures On Wavelets"*, CBMS-NSF Regional conference series in applied mathematics (1992).
[6] F. De Marie, H. G. Feichtinger, K. Nowak, *Uniform eigenvalue estimates for time-frequency localization operators,* J. London Math. Soc. (2), 65 (2002), 720–732.
[7] D. L. Donoho, P. B. Stark, *Uncertainty principles and signal recovery*, SIAM J. Appl. Math., 49 (1989), 906-931.
[8] H. G. Feichtinger, K. Nowak, *A Szegö-type theorem for Gabor-Toeplitz localization operators*, Michigan Math. J., 49 (2001), 13-21.
[9] K. Gröchenig, J. Toft, *The range of localization operators and lifting theorems for modulation and Bargmann-Fock spaces*,Trans. Amer. Math. Soc., 365 (2013), 4475–4496.
[10] A. C. Hansen, *On the solvability complexity index, the n-pseudospectrum and approximations of spectra of operators*, J. Amer. Math. Soc. 24 (2011), 81-124.

1.5. Organization of the paper. We have organized the ideas in the following way. The next section describes the setting of multivariated functions which are bandlimited to a measurable set and the main notations, the main result and corresponding proof are presented. In the last section we outline the setting of Gabor localization operators and formulate an analogue of the quantitative Nyquist densities in the Gabor setting. The proofs are very similar to those in section 2 and are omitted.

2. Notations and main results

2.1. Time- and band- limiting operators. We will use the notation

$$Tf(\xi) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(t) e^{-i\xi t} dt$$

for the Fourier transform of a function $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$. The subspaces of $L^2(\mathbb{R}^d)$ consisting, respectively, of the functions supported in T and of those whose Fourier transform is supported in Ω are

$$\begin{aligned} \mathcal{D}(T) &= \{f \in L^2(\mathbb{R}^d) : f(x) = 0, x \notin T\} \\ \mathcal{B}(\Omega) &= \{f \in L^2(\mathbb{R}^d) : Tf(\xi) = 0, \xi \notin \Omega\}. \end{aligned}$$

Let D_T be the orthogonal projection of $L^2(\mathbb{R}^d)$ onto $\mathcal{D}(T)$, given explicitly by the multiplication of a characteristic function of the set T by f:

$$D_T f(t) = \chi_T(t) f(t)$$

and let B_Ω be the orthogonal projection of $L^2(\mathbb{R}^d)$ onto $\mathcal{B}(\Omega)$, given explicitly as

$$B_\Omega f = T^{-1}\chi_\Omega T f = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} h(x-y) f(y) dy,$$

where $Th = \chi_\Omega$. The following proposition, comprising Lemma 1 and Theorem 1 of [14] gives important information concerning the spectral problem associated to the operator $D_{rT}B_\Omega D_{rT}$. This information will be essential in our proofs.

Proposition 1. [14] *The operator $D_{rT}B_\Omega D_{rT}$ is bounded by* 1*, self-adjoint, positive, and completely continuous. Denoting its set of eigenvalues, arranged*

in nonincreasing order, by $\{\lambda_k(r,T,\Omega)\}$*, we have*

$$\begin{aligned}\sum_{k=0}^{\infty}\lambda_k(r,T,\Omega) &= r^d\,(2\pi)^{-d}\,|T|\,|\Omega|\\ \sum_{k=0}^{\infty}\lambda_k^2(r,T,\Omega) &= r^d\,(2\pi)^{-d}\,|T|\,|\Omega| - o(r^d).\end{aligned}$$

Moreover, given $0<\gamma<1$*, the number* $M_r(\gamma)$ *of eigenvalues which are not smaller than* γ*, satisfies, as* $r\to\infty$*,*

$$M_r(\gamma) = (2\pi)^{-d}\,|T|\,|\Omega|\,r^d + o(r^d).$$

We are now in a position to state and prove our main theorem. The lower inequality is proved by constructing a set of orthonormal functions of $L^2(\mathbb{R}^d)$ satisfying (1.6). The proof of the upper inequality uses some of the techniques contained in Landau's proof of the non-hermitian Szegö-type theorem [14, Theorem 3].

Theorem 1. *Let* $\eta(\epsilon, rT, \Omega)$ *stand for the maximum number of orthonormal functions of* $L^2(\mathbb{R}^d)$ *such that*

$$\|D_{rT}B_\Omega D_{rT}f - f\| \le \epsilon. \tag{2.1}$$

Then, as $r\to\infty$*, the following inequalities hold:*

$$\frac{|T|\,|\Omega|}{(2\pi)^d}(1+\epsilon^2) \le \lim_{r\to\infty}\frac{\eta(\epsilon,rT,\Omega)}{r^d} \le \frac{|T|\,|\Omega|}{(2\pi)^d}\left(1-4\epsilon^2\right)^{-1}. \tag{2.2}$$

Proof: We first prove the lower inequality in (2.2). Let $\sigma>0$ and $\mathcal{F}=\{\phi_k\}$ be the normalized system of eigenfunctions of the operator $D_{rT}B_\Omega D_{rT}$ with eigenvalues $\lambda_k > 1-\sigma$. Given $f\in L^2(\mathbb{R}^d)$, write

$$f = \sum a_k\phi_k + h, \tag{2.3}$$

with $h\in Ker\,(D_{rT}B_\Omega D_{rT})$. Then

$$D_{rT}B_\Omega D_{rT}f = \sum a_k\lambda_k\phi_k \tag{2.4}$$

and

$$\begin{aligned}\|D_{rT}B_\Omega D_{rT}f - f\|^2 &= \left\|\sum(1-\lambda_j)a_k\phi_k + h\right\|^2\\ &\le \sigma\sum|a_k|^2 + \|h\|^2\\ &= \sigma^2\|f\|^2 + (1-\sigma^2)\|h\|^2.\end{aligned} \tag{2.5}$$

for all $f,g\in L^2(\mathbb{R}^d)$. These operators have been introduced in time-frequency analysis by Daubechies [4]. Since then, applications and connections to several mathematical topics, namely complex and harmonic analysis [17], [2], [3], [9] have been found. The eigenvalue problem has been object of a detailed study in [16], [8] and [6].

The image of $L^2(\mathbb{R}^d)$ under the Gabor transform with the window g will be named as the *Gabor space* $\mathcal{G}_g$. It is the following subspace of $L^2(\mathbb{R}^{2d})$:

$$\mathcal{G}_g = \left\{V_g f : f\in L^2(\mathbb{R}^d)\right\}.$$

The reproducing kernel of the Gabor space $\mathcal{G}_g$ is

$$K_g(z,w) = \langle \pi_z g, \pi_w g\rangle_{L^2(\mathbb{R}^d)} \tag{3.2}$$

and the projection operator $\mathcal{P}_g : L^2(\mathbb{R}^{2d})\to\mathcal{G}_g$,

$$\mathcal{P}_g F(z) = \int F(w)\overline{K_g(z,w)}dw.$$

It is shown in [16] that, for $F\in\mathcal{G}_g$,

$$\mathcal{V}_g\mathcal{C}_S\mathcal{V}_g^{-1}F(z) = \int_S F(w)\overline{K_g(z,w)}dw = \mathcal{P}_g D_S F(z).$$

For the whole $L^2(\mathbb{R}^{2d})$ one can write

$$\mathcal{V}_g\mathcal{C}_S\mathcal{V}_g^{*} = \mathcal{P}_g D_S.$$

Thus, the spectral properties of $\mathcal{C}_S$ are identical to those of $\mathcal{P}_g D_S$. Moreover, the operator $D_S\mathcal{P}_g D_S$ in $L^2(\mathbb{R}^{2d})$ and the operator $\mathcal{P}_g D_S$ have the same nonzero eigenvalues with multiplicity (see Lemma 1 in [16]). The analogue of Proposition 1 in this context is the following.

Proposition 2. [16] *The operator* $D_{rS}\mathcal{P}_g D_{rS}$ *is bounded by* 1*, self-adjoint, positive, and completely continuous. Denoting its set of eigenvalues, arranged in nonincreasing order, by* $\{\lambda_k(rS)\}$*, we have*

$$\begin{aligned}\sum_{k=0}^{\infty}\lambda_k(rS) &= r^d\,|S|\\ \sum_{k=0}^{\infty}\lambda_k^2(rS) &= r^d\,|S| - o(r^d).\end{aligned}$$

As a result of the previous paragraph, after excluding from $\mathcal{N}$ at most $\sum \frac{\dim \mathcal{E}(\delta)}{\delta}$ elements, those remaining are in $\mathcal{F}(\delta)$. Since, from Proposition 1, we have $\dim \mathcal{E}(\delta) = o(r^d)$, there are $\eta(\epsilon, rT, \Omega) - o(r^d)$ functions in $\mathcal{N} \cap \mathcal{F}(\delta)$. Let f be one of them. Now we can use (2.1), (2.13) and the triangle inequality to obtain

$$1 - \|B_\Omega D_{rT} f\|^2 \leq \|B_\Omega D_{rT} f - f\| \leq 2\varepsilon,$$

leading to $\|B_\Omega D_{rT} f\|^2 \geq 1 - 4\varepsilon^2$, for each of the $\eta(\epsilon, rT, \Omega) - o(r^d)$ orthonormal functions. Since $\|B_\Omega D_{rT} f\|^2 = \langle D_{rT} B_\Omega D_{rT} f, f \rangle$, the sum of these terms for any orthonormal set cannot exceed the trace of $D_{rT} B_\Omega D_{rT}$. Thus, using the trace obtained in Proposition 1, we conclude that

$$(1 - 4\varepsilon^2) \left(\eta(\epsilon, rT, \Omega) - o(r^d) \right) \leq \sum_{k=0}^{\infty} \lambda_k(r, T, \Omega) = r^d (2\pi)^{-d} |T| |\Omega|,$$

leading to the upper inequality in (2.2). ∎

Remark 1. *In the case where T and Ω are finite unions of bounded intervals, the term $o(r)$ in Proposition 1 can be replaced by* $\log r$ [15], [13]. *Thus, (1.4) follows using this estimate in our proofs of Theorem 1 and Theorem 2. See the recent monograph* [12] *for more estimates on the eigenvalues of the time- and band- limiting operator.*

Remark 2. *It is possible to obtain analogues of Theorem 1 and Theorem 2 in the set up of the Hankel transform. The result corresponding to Proposition 1 has been proved in* [1].

3. Gabor localization operators

The Gabor (or short-time Fourier) transform of a function or distribution f with respect to a window function $g \in L^2(\mathbb{R}^d)$ is defined to be, for $z = (x, \xi) \in \mathbb{R}^{2d}$:

$$\mathcal{V}_g f(z) = \mathcal{V}_g f(x, \xi) = \int_{\mathbb{R}^d} f(t) \overline{g(t - x)} e^{-2\pi i \xi t} dt. \tag{3.1}$$

The localization operator which concentrates the time-frequency content of a function in the region S operator $\mathcal{C}_S : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ can be defined weakly as

$$\langle \mathcal{C}_S f, h \rangle = \int \int_S \mathcal{V}_g f(x, \xi) \overline{\mathcal{V}_g h(x, \xi)} dx d\xi,$$

Now let n be an integer number such that

$$n \leq \frac{1}{\varepsilon^2} \leq n + 1 \tag{2.6}$$

and consider the following partition of $\mathcal{F}$ into subsets $\mathcal{F}_i$, each of them containing n functions:

$$\mathcal{F} = \mathcal{F}_1 \cup ... \cup \mathcal{F}_l \cup \mathcal{F}_{residual}, \tag{2.7}$$

where the partition is made in such a way that the set $\mathcal{F}_{residual}$ contains only $o(r)$ functions. This is possible to do because Proposition 1 tells us that $\#\mathcal{F} = \frac{|T||\Omega|}{(2\pi)^d} + o(r)$. For each set $\mathcal{F}_i$ choose h_i such that $h_i \in Ker\,(D_{rT} B_\Omega D_{rT})$ and such that

$$\langle h_i, h_j \rangle = \delta_{i,j}. \tag{2.8}$$

This can be done since $Ker\,(D_{rT} B_\Omega D_{rT})$ has infinite dimension, due to the inclusion $\mathcal{D}(\mathbb{R}^d - rT) \subset Ker\,(D_{rT} B_\Omega D_{rT})$. Now, for each i, let $\{\psi_j^{(i)}\}_{j=1}^{n+1}$ be a set of linear combination of functions of $\mathcal{F}$ such that

$$\left\langle \psi_k^{(i)}, \psi_j^{(i)} \right\rangle = \begin{cases} -\frac{1}{n+1} & if \quad k \neq j \\ 1 - \frac{1}{n+1} & if \quad k = j \end{cases}, \tag{2.9}$$

which can be constructed using a linear algebra argument as in the next paragraph.

Consider a linear transformation $T : \mathbb{R}^n \longrightarrow \mathcal{F}_i$ mapping each vector of the canonical basis of $\mathbb{R}^n$ to each of the given n orthogonal functions of $\mathcal{F}_i$. Let V be the subspace of $\mathbb{R}^{n+1}$ which is orthogonal to the vector $v_0 = \left[\sqrt{\frac{1}{n+1}}, ..., \sqrt{\frac{1}{n+1}}\right]^T \in \mathbb{R}^{n+1}$ and let $\{v_1, ..., v_n\}$ be an orthonormal basis of V. Clearly, $\|v_0\| = 1$ and, for $i = 1, ..., n$, $\langle v_0, v_i \rangle = 0$. Thus, the matrix

$$Q = \begin{bmatrix} v_0 & v_1 & ... & v_{n+1} \end{bmatrix} \in \mathbb{R}^{(n+1)\times(n+1)}$$

is orthogonal. If $u_1, ..., u_{n+1} \in \mathbb{R}^{n+1}$ are the rows of Q then

$$Q^T = \begin{bmatrix} u_1 & ... & u_{n+1} \end{bmatrix} \in \mathbb{R}^{(n+1)\times(n+1)}$$

is also orthogonal, we have $\langle u_i, u_j \rangle = \delta_{i,j}$. Let $u'_1, ..., u'_{n+1} \in \mathbb{R}^n$ be the rows of Q without the elements of the first column. They satisfy

$$\langle u'_k, u'_j \rangle = \langle u_k, u_j \rangle - \frac{1}{n+1} = \begin{cases} -\frac{1}{n+1} & if \quad k \neq j \\ 1 - \frac{1}{n+1} & if \quad k = j \end{cases},$$

and the functions in (2.9) are obtained setting $\psi_j^{(i)} = T u'_j$.

We are now in a position to construct the desired orthonormal system. Define a sequence of orthonormal functions $\{\Phi_j^{(i)}\}_{i=1}^{l}$ using the functions $\psi_j^{(i)}$ from (2.9):

$$\Phi_j^{(i)} = \psi_j^{(i)} + \sqrt{\frac{1}{n+1}} h_i. \tag{2.10}$$

Since $\psi_j^{(i)}$ are linear combinations of elements of $\mathcal{F} = \{\phi_k\}$, (2.10) is a representation of the form (2.3). Thus, (2.9) and (2.8) show that indeed $\left\langle \Phi_k^{(i)}, \Phi_j^{(i)} \right\rangle = \delta_{k,j}$ and we can apply (2.5) and (2.6) to obtain

$$\left\| D_{rT} B_\Omega D_{rT} \Phi_j^{(i)} - \Phi_j^{(i)} \right\|^2 \leq \sigma^2 + (1-\sigma^2)\varepsilon^2. \tag{2.11}$$

By construction, $\#\{\Phi_j^{(i)}\}_{j=1}^{n+1} = n+1$ and $\#\mathcal{F}_i = n$. Thus,

$$\#\{\Phi_j^{(i)}\}_{j=1}^{n+1} = \frac{n+1}{n} \#\mathcal{F}_i.$$

Now, the cardinality of the union of all the sequences $\{\Phi_j^{(i)}\}$ obtained according to the above procedure is

$$\begin{aligned}
\# \left[\cup_{i=1}^{l} \{\Phi_j^{(i)}\}_{j=1}^{n+1} \right] &= \frac{n+1}{n} \# \left[\cup_{i=1}^{l} \mathcal{F}_i \right] \\
&= \frac{n+1}{n} \# \left[\mathcal{F} - \mathcal{F}_{residual} \right] \\
&= \frac{n+1}{n} (r^d (2\pi)^{-d} |T| |\Omega| + o(r)) \\
&\geq \frac{\frac{1}{\varepsilon^2}+1}{\frac{1}{\varepsilon^2}} r^d (2\pi)^{-d} |T| |\Omega| + o(r) \\
&= (1+\varepsilon^2) r^d (2\pi)^{-d} |T| |\Omega| + o(r).
\end{aligned}$$

We have used Proposition 1 in the third equality (the fact that the dimension of $\mathcal{F}$ is $r^d (2\pi)^{-d} |T| |\Omega| + o(r)$ and the fact that $\mathcal{F}_{residual}$ contains only $o(r)$ functions). Since the resulting estimate

$$\# \left[\cup_{i=1}^{l} \{\Phi_j^{(i)}\}_{j=1}^{n+1} \right] \geq (1+\varepsilon^2) r^d (2\pi)^{-d} |T| |\Omega| + o(r)$$

holds for any $\sigma > 0$, we can take the limit $\sigma \to 0$ in (2.11) to yield

$$\left\| D_{rT} B_\Omega D_{rT} \Phi_j^{(i)} - \Phi_j^{(i)} \right\| \leq \varepsilon,$$

for every i and j. Thus, $\cup_{i=1}^{l} \{\Phi_j^{(i)}\}_{j=1}^{n+1}$ contains at least $(1+\varepsilon^2) r^d (2\pi)^{-d} |T| |\Omega| + o(r)$ orthonormal functions satisfying (2.1). This proves the lower inequality in (2.2).

Let us now prove the upper inequality in (2.2). Consider again $f = \sum a_k \phi_k + h$ with $h \in Ker\,(D_{rT} B_\Omega D_{rT})$. Then, using (2.4) and

$$\|B_\Omega D_{rT} f\|^2 = \langle D_{rT} B_\Omega D_{rT} f, f \rangle = \sum |a_k|^2 \lambda_k,$$

together with the fact that D_{rT} is a projection, one can write

$$\|B_\Omega D_{rT} f - D_{rT} B_\Omega D_{rT} f\|^2 = \|B_\Omega D_{rT} f\|^2 - \|D_{rT} B_\Omega D_{rT} f\|^2 = \sum |a_k|^2 \lambda_k (1-\lambda_k). \tag{2.12}$$

Now, for $\delta > 0$ define $\mathcal{E}(\delta)$ as the subspace generated by the eigenfunctions of $D_{rT} B_\Omega D_{rT}$ such that the corresponding eigenvalues satisfy $\delta < \lambda_k < 1-\delta$, and let

$$\mathcal{F}(\delta) = \left\{ f \in L^2(\mathbb{R}^d) : \|f\| = 1 \quad \sum_{\delta < \lambda_k < 1-\delta} |a_k|^2 \leq \delta \right\}.$$

For $f \in \mathcal{F}(\delta)$,

$$\begin{aligned}
&\|B_\Omega D_{rT} f - D_{rT} B_\Omega D_{rT} f\|^2 \\
&= \sum_{\lambda_k \leq \delta} |a_k|^2 \lambda_k (1-\lambda_k) + \sum_{\delta < \lambda_k < 1-\delta} |a_k|^2 \lambda_k (1-\lambda_k) + \sum_{\lambda_k \geq 1-\delta} |a_k|^2 \lambda_k (1-\lambda_k) \leq 2\delta.
\end{aligned}$$

Thus, δ can be choosen in such a way that

$$\|B_\Omega D_{rT} f - D_{rT} B_\Omega D_{rT} f\| \leq \varepsilon. \tag{2.13}$$

Let us assume the existence of a set $\mathcal{N}$ of $\eta(\epsilon, rT, \Omega)$ orthonormal functions of $L^2(\mathbb{R}^d)$ satisfying (2.1). To estimate how many of them belong to $\mathcal{F}(\delta)$, consider two subspaces $\mathcal{E}$ and $\mathcal{G}$ with corresponding projections E, G, and dimensions e and g respectively, with $e < g$. Let $v_1, ..., v_g$ be an orthonormal set in $\mathcal{G}$. Then $\sum \|Ev_i\|^2 = \sum (Ev_i, v_i) = \sum (GEGv_i, v_i)$ represents the trace of the operator GEG, independent of the choice of basis. Choose the basis $\{w_i\}$ such that the first vectors are in $\mathcal{GE}$ and the remaining vectors in the orthogonal complement in $\mathcal{G}$ of $\mathcal{GE}$. For each of the latter, $(GEGw, w) = 0$, while the dimension of $\mathcal{GE}$ is at most e. Hence $\sum \|Ev_i\|^2 = \sum_1^g (Ew_i, w_i) \leq \sum_1^e (GEGw_i, w_i) \leq e$. Thus, the number of orthonormal vectors $\{v_i\}$ for which $\sum \|Ev_i\|^2 \geq \delta$ cannot exceed e/δ.