Subspaces with extra invariance nearest to observed data [☆]

C. Cabrelli [a,b,*], C.A. Mosquera [a,b]

[a] *Departamento de Matemática, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, Pabellón I, 1428 Buenos Aires, Argentina*
[b] *IMAS-CONICET, Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina*

A R T I C L E   I N F O

A B S T R A C T

Given an arbitrary finite set of data $\mathcal{F} = \{f_1, \ldots, f_m\} \subset L^2(\mathbb{R}^d)$ we prove the existence and show how to construct a "small shift invariant space" that is "closest" to the data $\mathcal{F}$ over certain class of closed subspaces of $L^2(\mathbb{R}^d)$. The approximating subspace is required to have extra-invariance properties, that is to be invariant under translations by a prefixed additive subgroup of $\mathbb{R}^d$ containing $\mathbb{Z}^d$. This is important for example in situations where we need to deal with jitter error of the data. Here small means that our solution subspace should be generated by the integer translates of a small number of generators. An expression for the error in terms of the data is provided and we construct a Parseval frame for the optimal space.

We also consider the problem of approximating $\mathcal{F}$ from generalized Paley–Wiener spaces of $\mathbb{R}^d$ that are generated by the integer translates of a finite number of functions. That is finitely generated shift invariant spaces that are translation invariant. We characterize these spaces in terms of multi-tile sets of $\mathbb{R}^d$, and show the connections with recent results on Riesz basis of exponentials on bounded sets of $\mathbb{R}^d$. Finally we study the discrete case for our approximation problem.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

Let $\mathcal{H}$ be a Hilbert space, $\mathcal{C}$ a class of closed subspaces of $\mathcal{H}$ and $\mathcal{F} = \{f_1, \ldots, f_m\}$ a finite set of elements in $\mathcal{H}$.

In this article we study the existence and show how to construct an optimal subspace $\mathcal{S}$ in the class $\mathcal{C}$ that minimizes the distance to the given data $\mathcal{F}$, in the sense that $\mathcal{S}$ minimizes the functional $\pi(\mathcal{F}, \mathcal{S})$ over $\mathcal{C}$. The functional is defined as

[*] Corresponding author at: Departamento de Matemática, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, Pabellón I, 1428 Buenos Aires, Argentina. Fax: +54 11 4788 1824.
E-mail addresses: cabrelli@dm.uba.ar (C. Cabrelli), mosquera@dm.uba.ar (C.A. Mosquera).

$$\pi(\mathcal{F}, \mathcal{S}) = \sum_{j=1}^{m} \|f_j - P_{\mathcal{S}} f_j\|^2, \tag{1}$$

where $P_{\mathcal{S}}$ denotes the orthogonal projection on the subspace $\mathcal{S}$.

The motivation to find an optimal subspace in $\mathcal{C}$ is, that in many situations one wants to choose a model for a certain class of data. Instead of imposing some conditions on the data to fit some known model, the idea is to define a large class of subspaces convenient for the application at hand, and find from there the one that "best fits" the data under study.

The signals that need to be modeled are ideally low dimensional but living in a high dimensional space. However, since in applications they are often corrupted by noise, they become high dimensional, however they are close to a low dimensional subspace, which is the space one seeks.

When the Hilbert space is $L^2(\mathbb{R}^d)$ it is natural to consider as a model for our data the class of shift invariant spaces (SIS), that is, closed subspaces of $L^2(\mathbb{R}^d)$ that are invariant under translations by integers. These spaces have been used in approximation theory, harmonic analysis, wavelet theory, sampling theory and signal processing (see, e.g., [5,15,18,20] and the references therein).

Often, in applications, it is assumed that the signals under study belong to some shift invariant space V generated by the translations of a finite set of functions $\Phi = \{\varphi_1, \ldots, \varphi_m\}$, i.e., $V = S(\Phi) = \overline{\mathrm{span}}\{T_k \varphi_i : k \in \mathbb{Z}^d, i = 1, \ldots, m\}$. The choice of the particular finitely generated shift invariant space typically is not deduced from a set of signals. For example in sampling theory, a classical assumption is that the signals to be sampled are band-limited, that is, they belong to the shift invariant space V generated by $\varphi(x) = \mathrm{sinc}(x)$. However, the band-limited assumption is not very realistic in many applications. Thus, it is natural to search for a finitely generated shift invariant space that is nearest to a set of some observed data.

In this paper we study the case when $\mathcal{H} = L^2(\mathbb{R}^d)$, and the approximating subspaces are shift invariant spaces that have extra-invariance, that is: If M is a subgroup of $\mathbb{R}^d$ such that $\mathbb{Z}^d \subset M$ we will say that $S(\varphi_1, \ldots, \varphi_m)$ is M extra-invariant if

$$\overline{\mathrm{span}}\{T_k \varphi_j : j = 1, \ldots, m, \ k \in \mathbb{Z}^d\} = \overline{\mathrm{span}}\{T_\alpha \varphi_j : j = 1, \ldots, m, \ \alpha \in M\}.$$

Therefore, the space $S(\varphi_1, \ldots, \varphi_m)$ is invariant under translates other than the integers, even though it is generated by the integer translates of a finite set of functions. Such spaces with extra-invariance are important in applications specially in those where the jitter error is an issue.

We first consider the case when the subgroup M is a proper subgroup of $\mathbb{R}^d$ that contains $\mathbb{Z}^d$. For that case we obtain one of the main contributions of this paper. We prove that for any finite set of data $\mathcal{F} = \{f_1, \ldots, f_m\} \subset L^2(\mathbb{R}^d)$, for any proper subgroup M containing $\mathbb{Z}^d$ and for any $\ell \in \mathbb{N}$ there always exists a SIS V of length at most ℓ with extra-invariance M whose distance (in the sense of (1)) to the data $\mathcal{F}$ is the smallest possible among all the SIS of length smaller or equal than ℓ that are M extra-invariant. (Here, the length of a SIS is the cardinal of the smallest set of generators.)

We construct a solution V and provide a set of generators whose integer translates form a tight frame of V. An expression for the exact value of the error $\mathcal{E}(\mathcal{F}, V)$ between the data and the optimal subspace is also obtained using the eigenvalues of some special matrix.

Next, we consider the approximation problem for the class of generalized Paley–Wiener spaces. Given a measurable set $\Omega \subset \mathbb{R}^d$ (not necessarily bounded), the generalized Paley–Wiener space PW_Ω associate to Ω is the subspace of $L^2(\mathbb{R}^d)$ corresponding to the functions whose Fourier transform vanished outside Ω. A generalized Paley–Wiener space is always invariant under translations by the whole group $\mathbb{R}^d$. In particular is a SIS of $L^2(\mathbb{R}^d)$ that not necessarily need to be finitely generated. Under the hypothesis that PW_Ω has a Riesz basis of integer translates, we proved that PW_Ω is finitely generated if and only if Ω is a multi-tile (see Proposition 4.3). That is Ω is a multi-tile if and only if PW_Ω has extra-invariance $M = \mathbb{R}^d$.

We study our approximation problem for those generalized Paley–Wiener spaces. We describe for this case how to construct a set of generators and show an interesting connection with recent results about bases of exponentials. This completes all the cases of extra-invariance when M is not a proper subgroup.

Finally we consider a similar problem when the Hilbert space is $\ell^2(\mathbb{Z}^d)$ and $\mathcal{C}$ is a conveniently chosen class of subspaces. We obtain for this case equivalent results to the ones for $L^2(\mathbb{R}^d)$. The approximation problem for the discrete case is related with the continuous case in a very interesting way that is described in Section 5.

1.1. Previous work

Let us now mention some previous related work. The problem of approximation of a set of data by shift invariant spaces (without the extra-invariance restriction) started in [2] where the authors proved the existence of a minimizer for (1) over the class of low dimensional subspaces in a Hilbert space $\mathcal{H}$ and also over the class of shift invariant spaces in $L^2(\mathbb{R}^d)$.

In [4] the case of multiple subspaces was considered in the finitely dimensional case. That is, the authors found a union of low dimensional subspaces that best fits a given set of data in $\mathbb{R}^d$ and provided an algorithm to find it. In 2011, using dimensional reduction techniques, this algorithm was improved (see [1]).

Further, in [7] the authors found necessary and sufficient conditions for the existence of optimal subspaces in the general context of Hilbert spaces. However they did not provide a way to construct them.

The first result for approximation of a finite set of data using shift invariant spaces with extra-invariance constrains appears in [6], where the authors consider principal shift invariant spaces in one variable and they assume that the space has a generator with orthogonal integer translates, which is a key element in their proof. So the techniques of this particular case do not apply to our general case.

1.2. Organization of the paper

The paper is organized as follows. In Section 2 we set the definitions and results that we need about shift invariant spaces, extra-invariance and the approximation problem for the case of shift invariant spaces in $L^2(\mathbb{R}^d)$.

The main results of the paper are stated and proved in Sections 3, 4 and 5. In Section 3 we present the M extra-invariant case for shift invariant spaces, in Section 4 the case of Paley–Wiener spaces and finally we consider a discrete case, in Section 5.

2. Preliminaries

We begin with a review of the basic results and definitions that will be needed in subsequent sections. The known results are generally stated without proofs, but we provide references where the proofs can be found. Also, we introduce some of our notational conventions. For the definitions of Riesz bases and frames in Hilbert spaces we refer the reader to [11,16] and the references therein.

2.1. Shift invariant spaces

The structure of these spaces has been deeply analyzed (see for example [9,12,13,17,21]).

Definition 2.1. A closed subspace $V \subset L^2(\mathbb{R}^d)$ is said to be a *shift invariant space* if

$$f \in V \implies T_k f \in V, \quad \text{for any} \ \ k \in \mathbb{Z}^d,$$

where T_k is the translation by the vector $k \in \mathbb{Z}^d$, i.e. $T_k f(x) = f(x - k)$.

For any subset $\Phi \subset L^2(\mathbb{R}^d)$ we define

$$S(\Phi) = \overline{\text{span}}\{T_k\varphi \colon \varphi \in \Phi, k \in \mathbb{Z}^d\} \quad \text{and} \quad E(\Phi) = \{T_k\varphi \colon \varphi \in \Phi, k \in \mathbb{Z}^d\}.$$

We call $S(\Phi)$ the shift invariant space (SIS) generated by Φ. If $V = S(\Phi)$ for some finite set Φ we say that V is a *finitely generated* SIS, and a *principal* SIS if V can be generated by the integer translates of a single function.

For a finitely generated SIS $V \subset L^2(\mathbb{R}^d)$ we define the length of V as

$$\ell(V) = \min\{n \in \mathbb{N} \colon \exists \ \varphi_1, \ldots, \varphi_n \in V \text{ with } V = S(\varphi_1, \ldots, \varphi_n)\}.$$

In addition to the construction of a set of generators of the optimal space for the problems considered in this paper, it will be important to estimate the error of these approximations. In order to compute these errors we need to consider what is called the Gramian G_Φ for a family of functions $\Phi \subset L^2(\mathbb{R}^d)$.

More precisely, given $\Phi = \{\varphi_1, \ldots, \varphi_m\}$ a finite collection of functions in $L^2(\mathbb{R}^d)$, the *Gramian* G_Φ of Φ is the $m \times m$ matrix of $\mathbb{Z}^d$-periodic functions

$$[G_\Phi(\omega)]_{ij} = \sum_{k \in \mathbb{Z}^d} \widehat{\varphi}_i(\omega + k) \overline{\widehat{\varphi}_j(\omega + k)}. \tag{2}$$

The Gramian of Φ is determined a.e. by its values at any measurable set of representatives $\mathcal{U}$ of the quotient $\mathbb{R}^d/\mathbb{Z}^d$ and satisfies $G_\Phi(\omega)^* = G_\Phi(\omega)$ for a.e. $\omega \in \mathcal{U}$. We will take $\mathcal{U} = [-1/2, 1/2)^d$.

For a finitely generated SIS V, we can express the length of V in terms of the Gramian as follows (see [9,12,22])

$$\ell(V) = \operatorname*{essup}_{\omega \in \mathcal{U}}[\text{rk}(G_\Phi(\omega))] \tag{3}$$

where $\text{rk}(B)$ denotes the rank of a matrix B and Φ is a generator set for V.

One important property of the Gramian is given by the following lemma concerning the measurability of the eigenvalues and the existence of measurable eigenvectors of a non-negative matrix with measurable entries.

Lemma 2.2. *(See Lemma 2.3.5 of [21].) Let $G(\omega)$ be an $m \times m$ self-adjoint matrix of measurable functions defined on a measurable subset $E \subset \mathbb{R}^d$ with eigenvalues $\lambda_1(\omega) \geq \cdots \geq \lambda_m(\omega)$. Then the eigenvalues λ_i, $i = 1, \ldots, m$, are measurable functions on E and there exists an $m \times m$ matrix of measurable functions $U(\omega)$ on E such that $U(\omega)U^*(\omega) = I$ a.e. $\omega \in E$ and such that*

$$G(\omega) = U(\omega)\Lambda(\omega)U^*(\omega), \quad a.e. \ \omega \in E,$$

where $\Lambda(\omega) := diag(\lambda_1(\omega), \ldots, \lambda_m(\omega))$.

In [17], Helson introduced range functions and used this notion to completely characterize shift invariant spaces. Later on, several authors have used this framework to describe and characterize frames and bases of these spaces. See for example [9,10,12,13,21]. We will mention the required definitions and some known results that we need later for giving the proofs of our results. We refer to [9,12,13,21] for a complete description and the proofs.

Definition 2.3. Let $f \in L^2(\mathbb{R}^d)$ and fix $\mathcal{U} \subset \mathbb{R}^d$ to be a measurable set of representatives of the quotient $\mathbb{R}^d/\mathbb{Z}^d$. For $\omega \in \mathcal{U}$, the *fiber* $\tau f(\omega)$ of f at ω is the sequence

$$\tau f(\omega) = \{\widehat{f}(\omega + k)\}_{k \in \mathbb{Z}^d}.$$

Here $\widehat{f}$ denotes the Fourier transform of the function f, that is $\widehat{f}(\omega) = \int_{\mathbb{R}^d} e^{-2\pi i \omega x} f(x)\, dx$ when $f \in L^1(\mathbb{R}^d)$. We observe that if $f \in L^2(\mathbb{R}^d)$, then the fiber $\tau f(\omega)$ belongs to $\ell^2(\mathbb{Z}^d)$ for almost every $\omega \in \mathcal{U}$.

If V is a finitely generated SIS and $\omega \in \mathcal{U}$ we define the *fiber space* associated to V and ω as follows

$$J_V(\omega) = \overline{\{\tau f(\omega) : f \in V\}},$$

where the closure is taken in the norm of $\ell^2(\mathbb{Z}^d)$.

With the above definitions we have:

Lemma 2.4. *(See Proposition 5.6 of [8].) Let $V = S(\Phi)$ be a finitely generated SIS. Then*

$$\dim(J_V(\omega)) = rk\,(G_\Phi(\omega)), \quad a.e.\ \omega \in \mathcal{U}.$$

Lemma 2.5. *If $f \in L^2(\mathbb{R}^d)$, then*

(i) *the sequence $\tau f(\omega) = \{\widehat{f}(\omega + k)\}_{k \in \mathbb{Z}^d}$ is a well-defined sequence in $\ell^2(\mathbb{Z}^d)$ a.e. $\omega \in \mathcal{U}$.*
(ii) *$\|\tau f(\omega)\|_{\ell^2}$ is a measurable function of ω and*

$$\|f\|^2 = \|\widehat{f}\|^2 = \int_{\mathcal{U}} \|\tau f(\omega)\|_{\ell^2}^2\, d\omega.$$

Lemma 2.6. *Let V be a finitely generated SIS in $L^2(\mathbb{R}^d)$. Then we have*

(i) *$J_V(\omega)$ is a closed subspace of $\ell^2(\mathbb{Z}^d)$ for a.e. $\omega \in \mathcal{U}$.*
(ii) *$V = \{f \in L^2(\mathbb{R}^d) : \tau f(\omega) \in J_V(\omega)\ \text{for a.e. } \omega \in \mathcal{U}\}$.*
(iii) *For each $f \in L^2(\mathbb{R}^d)$ we have that $\|\tau(P_V f)(\omega)\|_{\ell^2}$ is a measurable function of the variable ω and*

$$\tau(P_V f)(\omega) = P_{J_V(\omega)}(\tau f(\omega)).$$

(iv) *Let $\varphi_1, \ldots, \varphi_m \in L^2(\mathbb{R}^d)$. We have that*
    (a) *$\{\varphi_1, \ldots, \varphi_m\}$ is a set of generators of V, if and only if $\{\tau\varphi_1(\omega), \ldots, \tau\varphi_m(\omega)\}$ spans $J_V(\omega)$ for a.e. $\omega \in \mathcal{U}$.*
    (b) *The integer translations of $\varphi_1, \ldots, \varphi_m$ are a frame (resp. Riesz basis) of V, if and only if $\tau\varphi_1(\omega), \ldots, \tau\varphi_m(\omega)$ are a frame (resp. Riesz basis) of $J_V(\omega)$ with the same frame (resp. Riesz) bounds, for a.e. $\omega \in \mathcal{U}$.*

2.2. Optimality for the class of SIS in $L^2(\mathbb{R}^d)$

In [2] the authors give a solution for the case where the approximation class is the class of SIS in $L^2(\mathbb{R}^d)$. For this, they reduce the optimization problem into an uncountable set of finite dimensional problems in the Hilbert space $\mathcal{H} = \ell^2(\mathbb{Z}^d)$.

Theorem 2.7. *(See Theorem 2.3 of [2].) Let $\mathcal{F} = \{f_1, \ldots, f_m\}$ be a set of functions in $L^2(\mathbb{R}^d)$. Let $\lambda_1(\omega) \geq \cdots \geq \lambda_m(\omega)$ be the eigenvalues of the Gramian $G_{\mathcal{F}}(\omega)$. Then, there exists $V^* \in \mathcal{V}^\ell = \{V : V\ \text{is a SIS of length at most } \ell\}$ such that*

$$\sum_{i=1}^{m} \|f_i - P_{V^*}f_i\|^2 \leq \sum_{i=1}^{m} \|f_i - P_V f_i\|^2, \quad \forall V \in \mathcal{V}^\ell.$$

Moreover, we have that

(1) *The eigenvalues $\lambda_i(\omega)$, $1 \leq i \leq m$ are $\mathbb{Z}^d$-periodic, measurable functions in $L^2(\mathcal{U})$ and the approximation error is given by,*

$$\mathcal{E}(\mathcal{F}, \ell) = \sum_{i=\ell+1}^{m} \int_{\mathcal{U}} \lambda_i(\omega)\, d\omega.$$

(2) *Let $\theta_i(\omega) = \lambda_i^{-1}(\omega)$ if $\lambda_i(\omega)$ is different from zero, and zero otherwise. Then, there exists a choice of measurable left eigenvectors $Y^1(\omega), \ldots, Y^\ell(\omega)$ with $Y^i = (y_1^i, \ldots, y_m^i)^t$, $i = 1, \ldots, \ell$, associated with the first ℓ largest eigenvalues of $G_{\mathcal{F}}(\omega)$ such that the functions defined by*

$$\widehat{\varphi}_i(\omega) = \theta_i(\omega) \sum_{j=1}^{m} y_j^i(\omega) \widehat{f}_j(\omega), \quad i = 1, \ldots, \ell, \; \omega \in \mathbb{R}^d$$

are in $L^2(\mathbb{R}^d)$. Furthermore, the corresponding set of functions $\Phi = \{\varphi_1, \ldots, \varphi_\ell\}$ is a generator set for the optimal subspace V^* and the set $\{\varphi_i(\cdot - k), k \in \mathbb{Z}^d, i = 1, \ldots, \ell\}$ is a Parseval frame for V^*.

2.3. Extra invariance

We will need some definitions and known results concerning extra-invariance for shift invariant spaces. These are described in this subsection.

Definition 2.8. Let $V \subset L^2(\mathbb{R}^d)$ be a SIS. We define the *invariance set* as follows

$$M := \{x \in \mathbb{R}^d \colon T_x f \in V, \forall f \in V\}.$$

In [3] (see also [8]), the authors proved that the invariance set of a shift invariance space $V \subset L^2(\mathbb{R}^d)$ is a closed additive subgroup of $\mathbb{R}^d$ that contains $\mathbb{Z}^d$. For instance, in the case of the line the invariant set of a shift invariant space could be $\mathbb{Z}, \frac{1}{n}\mathbb{Z}$ for some $n \in \mathbb{N}$ or $\mathbb{R}$.

Definition 2.9. Let $\Phi \subset L^2(\mathbb{R}^d)$. We will say that $V = S(\Phi)$ is *M extra-invariant* if $T_m f \in V$ for all $m \in M$ and for all $f \in V$.

If $M = \mathbb{R}^d$ then the space V is translation invariant but generated by the integer translates of the set Φ.

One example of a translation invariant space in $\mathbb{R}$ is the Paley–Wiener space of functions that are bandlimited to $[-1/2, 1/2]$ defined by

$$PW = \{f \in L^2(\mathbb{R}) \colon \text{supp}\,(\widehat{f}) \subset [-1/2, 1/2]\}.$$

It is easy to prove that for a measurable set $\Omega \subset \mathbb{R}^d$, the space

$$V_\Omega := \{f \in L^2(\mathbb{R}^d) \colon \text{supp}(\widehat{f}) \subset \Omega\} \tag{4}$$

is translation invariant. Moreover, Wiener's theorem (see [17]) proves that any closed translation invariant subspace of $L^2(\mathbb{R}^d)$ is of the form (4).

If V is a shift invariant space of length ℓ and M is an additive subgroup of $\mathbb{R}^d$ containing $\mathbb{Z}^d$, we will say that V has *extra-invariance* M if V is M-invariant. Note that in this case, if Φ is a set of generators of V, i.e. $V = S(\Phi)$, then

$$S(\Phi) = \overline{\text{span}}\{T_k\varphi \colon \varphi \in \Phi, k \in \mathbb{Z}^d\} = \overline{\text{span}}\{T_\alpha\varphi \colon \varphi \in \Phi, \alpha \in M\}.$$

In [3] the authors characterize those shift invariant spaces $V \subset L^2(\mathbb{R})$ that have extra-invariance. They show that either V is translation invariant, or there exists a maximum positive integer n such that V is $\frac{1}{n}\mathbb{Z}$-invariant.

The d-dimensional case is consider in [8]. There, a characterization of the extra-invariance of V when M is not all $\mathbb{R}^d$ is obtained. Given M a closed subgroup of $\mathbb{R}^d$ containing $\mathbb{Z}^d$ and $M^* = \{x \in \mathbb{R}^d \colon \langle x, m \rangle \in \mathbb{Z}$ $\forall m \in M\}$, the authors construct a special partition $\{B_\sigma\}_{\sigma \in \mathcal{N}}$ of $\mathbb{R}^d$, where each B_σ is an M^*-periodic set and the index set $\mathcal{N}$ is a section of the quotient $\mathbb{Z}^d/M^*$. More precisely, for each $\sigma \in \mathcal{N}$,

$$B_\sigma = \Omega + \sigma + M^* = \bigcup_{m^* \in M^*} (\Omega + \sigma) + m^*, \tag{5}$$

where Ω is a section of the quotient $\mathbb{R}^d/\mathbb{Z}^d$. We refer to [8] for more details.

Using this partition, for each $\sigma \in \mathcal{N}$, they define the subspaces associated to a given SIS V

$$V_\sigma = \{f \in L^2(\mathbb{R}^d) \colon \widehat{f} = \chi_{B_\sigma}\widehat{g}, \text{ with } g \in V\}. \tag{6}$$

Given $f \in L^2(\mathbb{R}^d)$ define for $\sigma \in \mathcal{N}$, the function f^σ by $\widehat{f^\sigma} = \widehat{f}\chi_{B_\sigma}$.

The authors give a characterization of the M-invariance of V in terms of the subspaces V_σ. More specifically they prove that

Theorem 2.10. *If $V \subset L^2(\mathbb{R}^d)$ is a SIS and M is a closed subgroup of $\mathbb{R}^d$ containing $\mathbb{Z}^d$, then the following are equivalent.*

(i) *V is M-invariant,*
(ii) *$V_\sigma \subset V$ for all $\sigma \in \mathcal{N}$,*
(iii) *$J_{V_\sigma}(\omega) \subset J_V(\omega)$ for almost every ω and each $\sigma \in \mathcal{N}$,*
(iv) *if $V = S(\Phi)$ then $\tau\varphi^\sigma(\omega) \in J_V(\omega)$ a.e. $\omega \in \mathcal{U}$ for all $\varphi \in \Phi$ and all $\sigma \in \mathcal{N}$.*

3. Optimality for the class of SIS with extra-invariance

Here we consider the approximation problem for the class of finitely generated SIS with extra-invariance under a given *proper* subgroup M of $\mathbb{R}^d$.

Let us start introducing some notation. Let $m, \ell \in \mathbb{N}$, M be a closed proper subgroup of $\mathbb{R}^d$ containing $\mathbb{Z}^d$ and $\mathcal{F} = \{f_1, \ldots, f_m\} \subset L^2(\mathbb{R}^d)$. Define

$$\mathcal{V}_M^\ell = \{V : V \text{ is a SIS of length at most } \ell \text{ and } V \text{ is } M\text{-invariant}\}. \tag{7}$$

Let $\mathcal{N} = \{\sigma_1, \ldots, \sigma_\kappa\}$ be a section of the quotient $\mathbb{Z}^d/M^*$ and $\{B_\sigma : \sigma \in \mathcal{N}\}$ the partition defined in (5).

For each $\sigma \in \mathcal{N}$, we consider $\mathcal{F}^\sigma = \{f_1^\sigma, \ldots, f_m^\sigma\} \subset L^2(\mathbb{R}^d)$ where, f_j^σ is such that $\widehat{f_j^\sigma} = \widehat{f_j}\chi_{B_\sigma}$ for $j = 1, \ldots, m$. Also, let $\widetilde{\mathcal{F}} = \{f_1^{\sigma_1}, \ldots, f_m^{\sigma_1}, \ldots, \ldots, f_1^{\sigma_\kappa}, \ldots, f_m^{\sigma_\kappa}\}$.

For each $\omega \in \mathcal{U}$ let $G_{\widetilde{\mathcal{F}}}(\omega)$ be the associated Gramian matrix of the vectors in $\widetilde{\mathcal{F}}$ with eigenvalues

$$\lambda_1(\omega) \geq \cdots \geq \lambda_{m\kappa}(\omega) \geq 0.$$

Using Lemma 2.2, we have that these eigenvalues are measurable functions.

Since $f_i^{\sigma_s}$ is orthogonal to $f_i^{\sigma_t}$ if $s \neq t$, the Gramian $G_{\widetilde{\mathcal{F}}}(\omega)$ is a diagonal block matrix with blocks $G_\sigma(\omega)$, $\sigma \in \mathcal{N}$. Here $G_\sigma(\omega)$ is the $m \times m$ Gramian associated to the data $\mathcal{F}^\sigma$. On the other hand, using Lemma 2.2 we have that

$$G_\sigma(\omega) = U_\sigma(\omega)\Lambda_\sigma(\omega)U_\sigma^*(\omega) \quad a.e. \ \omega \in \mathcal{U}$$

where U_σ are unitary and $\Lambda_\sigma(\omega) := \operatorname{diag}(\lambda_1^\sigma(\omega), \ldots, \lambda_m^\sigma(\omega)) \in \mathbb{C}^{m \times m}$ and they are also measurable matrices as in Lemma 2.2. We also have $\lambda_1^\sigma(\omega) \geq \cdots \geq \lambda_m^\sigma(\omega)$ for each $\sigma \in \mathcal{N}$.

Using the decompositions of the blocks G_σ we have that

$$G_{\widetilde{\mathcal{F}}}(\omega) = U(\omega)\Lambda(\omega)U^*(\omega) \tag{8}$$

where U has blocks U_σ in the diagonal, and Λ is diagonal with blocks Λ_σ. We want to recall here that for almost each ω the matrix $\Lambda(\omega)$ collects all the eigenvalues of the Gramian $G_{\widetilde{\mathcal{F}}}(\omega)$ and the columns of the matrix $U(\omega)$ are the associated left eigenvectors. Note that an eigenvector associated to the eigenvalue $\lambda_j^\sigma(\omega)$ has all the components not corresponding to the block σ equal to zero.

Now for each fixed $\omega \in \mathcal{U}$, we consider $\{(i_1(\omega), j_1(\omega)), \ldots, (i_n(\omega), j_n(\omega))\}$ with $i_s(\omega) \in \mathcal{N}$ and $j_s(\omega) \in \{1, \ldots, m\}$ and $n = m\kappa$ such that

$$\lambda_{j_1(\omega)}^{i_1(\omega)} \geq \cdots \geq \lambda_{j_n(\omega)}^{i_n(\omega)} \geq 0$$

are the ordered eigenvalues of $G_{\widetilde{\mathcal{F}}}(\omega)$, with corresponding left eigenvectors $Y^{(i_s(\omega), j_s(\omega))} \in \mathbb{C}^n$, for $s = 1, \ldots, n$.

Here $i_s(\omega)$ indicates the block of the matrix $G_{\widetilde{\mathcal{F}}}(\omega)$ in which the eigenvalue $\lambda_{j_s(\omega)}^{i_s(\omega)}(\omega)$ is found and $j_s(\omega)$ indicates the displacement in this block of the matrix $G_{\widetilde{\mathcal{F}}}(\omega)$. More precisely, we have that $\lambda_{j_s(\omega)}^{i_s(\omega)}(\omega)$ coincides with $\lambda_{(i_s(\omega)-1)m+j_s(\omega)}(\omega)$, the $((i_s(\omega)-1)m+j_s(\omega))$-th eigenvalue of $G_{\widetilde{\mathcal{F}}}(\omega)$. When $\omega \in \mathcal{U}$ is fixed, we will write i_s instead of $i_s(\omega)$ and j_s instead of $j_s(\omega)$.

We will prove now that $\gamma_s(\omega) := \lambda_{j_s(\omega)}^{i_s(\omega)}(\omega)$ is measurable as a function on ω for each $s = 1, \ldots, n$.

Let $s \in \{1, \ldots, n\}$ fixed. Let $i_s(\omega) \in \mathcal{N}$ and $j_s(\omega) \in \{1, \ldots, m\}$. We have that $\gamma_s(\omega) = \lambda_j^\sigma(\omega)$ for all $\omega \in E_{\sigma j} := \{\omega \in \mathcal{U} : i_s(\omega) = \sigma, j_s(\omega) = j\}$.

We observe that

$$E_{\sigma j} = \{\omega \in \mathcal{U} : i_s(\omega) = \sigma, j_s(\omega) = j\} = \{\omega \in \mathcal{U} : \lambda_s(\omega) = \lambda_j^\sigma(\omega)\}.$$

Using Lemma 2.2 applied to $G_{\widetilde{\mathcal{F}}}(\omega)$ and $G_\sigma(\omega)$, we have that λ_s and λ_j^σ are measurable functions of ω. Therefore $E_{\sigma j}$ are measurable sets.

We further observe that $\gamma_s(\omega) = \lambda_j^\sigma(\omega)$, for $\omega \in E_{\sigma j}$. So $\gamma_s(\omega)$ is a measurable function. A similar argument shows that the eigenvectors are measurable.

Finally we define $h_s : \mathbb{R}^d \to \mathbb{C}$, for $s = 1, \ldots, \ell$

$$h_s(\omega) := \theta_{j_s}^{i_s}(\omega) \sum_{k=1}^m y_{(i_s-1)m+k}^{(i_s, j_s)}(\omega)\widehat{f}_k^{i_s}(\omega), \tag{9}$$

where $\theta_{j_s}^{i_s}(\omega) = (\lambda_{j_s}^{i_s}(\omega))^{-1/2}$ if $\lambda_{j_s}^{i_s}(\omega) \neq 0$ and $\theta_{j_s}^{i_s}(\omega) = 0$ otherwise.

Now we are ready to state the main result of this section.

Theorem 3.1. *Let $m, \ell \in \mathbb{N}$, and M be a closed proper subgroup of $\mathbb{R}^d$ containing $\mathbb{Z}^d$. Assume that $\mathcal{F} = \{f_1, \ldots, f_m\} \subset L^2(\mathbb{R}^d)$ is given data and let $\mathcal{V}_M^\ell$ be the class defined in (7). Then, there exists a shift invariant space $V^* \in \mathcal{V}_M^\ell$ such that*

$$V^* = \underset{V \in \mathcal{V}_M^\ell}{argmin} \sum_{j=1}^m \|f_j - P_V f_j\|^2. \tag{10}$$

Furthermore, with the above notation,

(1) *The eigenvalues $\{\lambda_j^\sigma(\omega) : \sigma \in \mathcal{N}, j = 1, \ldots, m\}$, are $\mathbb{Z}^d$-periodic, measurable functions in $L^2(\mathcal{U})$ and the error of approximation is*

$$\mathcal{E}(\mathcal{F}, M, \ell) := \sum_{j=1}^m \|f_j - P_{V^*} f_j\|^2 = \int_\mathcal{U} \sum_{s=\ell+1}^{m\kappa} \lambda_{j_s}^{i_s}(\omega)\, d\omega.$$

(2) *The functions $\{h_1, \ldots, h_\ell\}$ defined in (9) are in $L^2(\mathbb{R}^d)$ and if $\varphi_1, \ldots, \varphi_\ell$ are defined by $\widehat{\varphi_j} = h_j$, then $\Phi = \{\varphi_1, \ldots, \varphi_\ell\}$ is a generator set for the optimal subspace V^* and the set $\{\varphi_i(\cdot - k), k \in \mathbb{Z}^d, i = 1, \ldots, \ell\}$ is a Parseval frame for V^*.*

Proof. Let $\mathcal{V}^\ell$ be the class defined in Theorem 2.7, that is $\mathcal{V}^\ell$ is the set of all shift invariant spaces V that can be generated by ℓ or less generators. (Note that we do not ask the elements of the class $\mathcal{V}^\ell$ to have extra-invariance.)

Define $V^* \in \mathcal{V}^\ell$ to be the optimal space given by Theorem 2.7 for the data $\widetilde{\mathcal{F}}$. That is,

$$\sum_{\sigma \in \mathcal{N}} \sum_{j=1}^m \|f_j^\sigma - P_{V^*} f_j^\sigma\|^2 \leq \sum_{\sigma \in \mathcal{N}} \sum_{j=1}^m \|f_j^\sigma - P_V f_j^\sigma\|^2 \qquad \forall\ V \in \mathcal{V}^\ell. \tag{11}$$

We claim that $V^* \in \mathcal{V}_M^\ell$ (in particular is M extra-invariant) and it is optimal in this class for the data $\mathcal{F}$, i.e.

$$\sum_{j=1}^m \|f_j - P_{V^*} f_j\|^2 \leq \sum_{j=1}^m \|f_j - P_V f_j\|^2 \qquad \forall\ V \in \mathcal{V}_M^\ell. \tag{12}$$

Let us prove first that V^* is M extra-invariant. For this we will check that the generators of V^* satisfy condition (iv) in Theorem 2.10. We have from (8) that the Gramian $G_{\widetilde{\mathcal{F}}}(\omega)$ can be decomposed as $G_{\widetilde{\mathcal{F}}}(\omega) = U(\omega)\Lambda(\omega)U^*(\omega)$ with eigenvalues $\{\lambda_j^\sigma(\omega) : \sigma \in \mathcal{N}, j = 1, \ldots, m\}$.

By Theorem 2.7, the ℓ generators of V^* have the form defined in (9),

$$\widehat{\varphi}_s(\omega) = \theta_{j_s}^{i_s}(\omega) \sum_{k=1}^m y_{(i_s-1)m+k}^{(i_s, j_s)}(\omega) \widehat{f_k^{i_s}}(\omega), \qquad \text{for } s = 1, \ldots, \ell. \tag{13}$$

From (13) it is clear that $\widehat{\varphi}_s$ is supported in B_{i_s}, since each $\widehat{f_k^{i_s}}$ is supported in B_{i_s}. Then if we apply the cut off operator to these generators we obtain $\widehat{\varphi}_s^\sigma(\omega) = \widehat{\varphi}_s(\omega)$ if $\sigma = i_s(\omega)$ and $\widehat{\varphi}_s^\sigma = 0$ otherwise. So, in any case $\varphi_s^\sigma \in V^*$ for all $\sigma \in \mathcal{N}$, $s = 1, \ldots, \ell$ which proves the M-invariance of V^*.

What is left now is to prove that V^* is optimal over the class $\mathcal{V}_M^\ell$, that is V^* satisfies equation (12). For this note that if $V \in \mathcal{V}_M^\ell$ then $V = \bigoplus_{\sigma \in \mathcal{N}} V_\sigma$. So we have for any $f \in L^2(\mathbb{R}^d)$,

$$\|P_V f\|^2 = \|P_V \sum_{\sigma \in \mathcal{N}} f^\sigma\|^2 = \|\sum_{\sigma \in \mathcal{N}} P_V f^\sigma\|^2 = \|\sum_{\sigma \in \mathcal{N}} P_{V^\sigma} f^\sigma\|^2 = \sum_{\sigma \in \mathcal{N}} \|P_{V^\sigma} f^\sigma\|^2 = \sum_{\sigma \in \mathcal{N}} \|P_V f^\sigma\|^2,$$

which implies together with (11) that

$$\sum_{j=1}^{m} \|P_{V^*} f_j\|^2 \geq \sum_{j=1}^{m} \|P_V f_j\|^2, \qquad \forall \ V \in \mathcal{V}_M^\ell.$$

The others claims of the theorem are a direct consequence of Theorem 2.7. $\quad\square$

4. Approximation with Paley–Wiener spaces

4.1. Preliminaries

In this section the class of approximation subspaces will be finitely generated SIS that are translation invariant. That is translation invariant spaces that are generated by the integer translates of a finite number of functions.

More precisely, given $\ell \in \mathbb{N}$ define $\mathcal{T}^\ell$ to be the set of all shift invariant spaces $V = S(\varphi_1, \ldots, \varphi_\ell)$ for some functions $\varphi_1, \ldots, \varphi_\ell$ in $L^2(\mathbb{R}^d)$, and such that V is translation invariant and the integer translates of $\{\varphi_1, \ldots, \varphi_\ell\}$ form a Riesz basis of V.

Given a set $\mathcal{F} = \{f_1, \ldots, f_m\} \subset L^2(\mathbb{R}^d)$, we want to find $V^* \in \mathcal{T}^\ell$ such that

$$V^* = \underset{V \in \mathcal{T}^\ell}{argmin} \sum_{j=1}^{m} \|f_j - P_V f_j\|^2. \tag{14}$$

Here P_V denotes the orthogonal projection on V.

Before going to the approximation problem, we will obtain a characterization of the class $\mathcal{T}^\ell$.

Using Wiener's theorem, we have that V is a translation invariant space in $L^2(\mathbb{R}^d)$ if and only if there exists a measurable set $\Omega \subset \mathbb{R}^d$ such that

$$V = \{f \in L^2(\mathbb{R}^d) : \hat{f}(\omega) = 0 \text{ a.e. } \omega \in \mathbb{R}^d \setminus \Omega\}.$$

Since Ω is unique up to measure zero, we will write $V = V_\Omega$.

Definition 4.1. Let $\Omega \subset \mathbb{R}^d$ be measurable and $L \subset \mathbb{R}^d$ be a countable set. We say that Ω *tiles* $\mathbb{R}^d$ *when translated by* L *at level* $\ell \in \mathbb{N}$ if

$$\sum_{t \in L} \chi_\Omega(\omega - t) = \ell, \quad \text{for a.e. } \omega \in \mathbb{R}^d.$$

In case of $L = \mathbb{Z}^d$ we will say that Ω is an ℓ *multi-tile.*

It is known (see for example [19]) that Ω is an ℓ multi-tile of $\mathbb{R}^d$, if and only if, up to measure zero, Ω is the union of ℓ measurable and disjoint 1 tile sets. i.e. Ω is a quasi-disjoint union of ℓ sets of representatives of $\mathbb{R}^d / \mathbb{Z}^d$.

Lemma 4.2. *A measurable set* $\Omega \subset \mathbb{R}^d$, ℓ *multi-tiles* $\mathbb{R}^d$ *if and only if*

$$\Omega = \Omega_1 \cup \cdots \cup \Omega_\ell \cup N,$$

where N *is a zero measure set, and the sets* Ω_j, $1 \leq j \leq \ell$ *are measurable, disjoint and each of them tiles* $\mathbb{R}^d$ *by translations on* $\mathbb{Z}^d$.

The following proposition characterizes the set Ω for the elements in $\mathcal{T}^\ell$.

Proposition 4.3. *A subspace V is in $\mathcal{T}^\ell$ if and only if $V = V_\Omega$ with Ω a measurable ℓ multi-tile of $\mathbb{R}^d$.*

Proof. Assume first that $V \in \mathcal{T}^\ell$, so $V = V_\Omega$ for some measurable $\Omega \subset \mathbb{R}^d$. Also, as a consequence of Wiener's theorem, for almost all $\omega \in \mathcal{U}$ we have $J_V(\omega) \cong \ell^2(O_\omega)$ with $O_\omega = \{k \in \mathbb{Z}^d : w + k \in \Omega\}$. To see this, we note that $J_V(\omega) \subset \ell^2(O_\omega)$. For the other inclusion, fix $\omega \in \mathcal{U}$. Using that $\Omega = \bigcup_{k \in \mathbb{Z}^d} E_k$ where $E_k = (\mathcal{U} + k) \cap \Omega$, we have that $k \in O_\omega$, if and only if $\omega + k \in E_k$. Hence, if $a \in \ell^2(O_\omega)$ consider the function $G_\omega(\xi) = \sum_{k \in O_\omega} a_k \chi_{E_k}(\xi)$. Since G_ω is in $L^2(\Omega)$, the function g defined by $\widehat{g} = G_\omega$ is in V, and $\widehat{g}(\omega + k) = a_k$ if $k \in O_\omega$. Therefore, $g \in V$ and $a = \tau g(\omega) \in J_V(\omega)$.

Now, since $V = S(\varphi_1, \dots, \varphi_\ell)$, and the integer translates of $\varphi_1, \dots, \varphi_\ell$ form a Riesz basis of V, using Lemma 2.6 we obtain that $\{\tau\varphi_1(\omega), \dots, \tau\varphi_\ell(\omega)\}$ form a Riesz basis of $J_V(\omega)$ with the same Riesz bounds for a.e. $\omega \in \mathcal{U}$. We conclude that $\dim(J_V(\omega)) = \ell$ a.e. $\omega \in \mathcal{U}$.

Since V is translation invariant, by the observation above $\dim(J_V(\omega)) = \#O_w$. Then $\#O_\omega = \ell$ for almost all $\omega \in \mathcal{U}$, which implies that Ω is an ℓ multi-tile. (Here $\#A$ denotes the cardinal of the set A.)

For the converse, assume that Ω is a measurable ℓ multi-tile of $\mathbb{R}^d$. Define $V = V_\Omega$. So, V is translation invariant.

By Lemma 4.2 we have that $\Omega = \Omega_1 \cup \cdots \cup \Omega_\ell$ up to a measure zero set, where each Ω_j is a set of representatives or $\mathbb{R}^d/\mathbb{Z}^d$. We define φ_j by its Fourier transform: $\widehat{\varphi}_j = \chi_{\Omega_j}$, $j = 1, \dots, \ell$.

Since $\{e^{2\pi i \omega k}\widehat{\varphi}_j : k \in \mathbb{Z}^d\}$ is an orthonormal basis of $L^2(\Omega_j)$, we have that $\{e^{2\pi i \omega k}\widehat{\varphi}_j : k \in \mathbb{Z}^d, j = 1, \dots, \ell\}$ is an orthonormal basis of $L^2(\Omega)$, and so, $\{t_k\varphi_j : k \in \mathbb{Z}^d, j = 1, \dots, \ell\}$ is an orthonormal basis of V, in particular a Riesz basis. $\square$

4.2. The approximation problem for Paley–Wiener spaces

Now we come back to our approximation problem. In order to find an optimal subspace in the class $\mathcal{T}^\ell$ for a set of data $\mathcal{F} = \{f_1, \dots, f_m\}$, it is enough to find the associated ℓ multi-tile Ω in $\mathbb{R}^d$. It is not difficult to see that if we allow Ω to be *any* ℓ multi-tile the minimum in (14) may not exist. So we will restrict Ω to be inside a cube that could be arbitrarily large. Let us fix $N \in \mathbb{N}$. Define

$$C_N := [-(N + 1/2), N + 1/2]^d,$$
$$M_N^\ell := \{\Omega \subset C_N : \Omega \text{ is measurable and } \ell \text{ multi-tiles } \mathbb{R}^d\} \text{ and}$$
$$\mathcal{T}_N^\ell := \{V \in \mathcal{T}^\ell : V = V_\Omega \text{ with } \Omega \in M_N^\ell\}.$$

With this notation we can state the main result of this section.

Theorem 4.4. *Assume that $m, \ell \in \mathbb{N}$ and a set $\mathcal{F} = \{f_1, \dots, f_m\} \subset L^2(\mathbb{R}^d)$, are given. Then for each $N \geq \ell$ there exists a Paley–Wiener space $V^* \in \mathcal{T}_N^\ell$ that satisfies*

$$V^* = \underset{V \in \mathcal{T}_N^\ell}{argmin} \sum_{j=1}^m \|f_j - P_V f_j\|^2, \tag{15}$$

where $\mathcal{T}_N^\ell$ is the class defined above.

Proof. First we observe that if a solution space V^* exists then

$$V^* = \underset{V \in \mathcal{T}_N^\ell}{argmin} \sum_{j=1}^m \|f_j - P_V f_j\|^2 = \underset{V \in \mathcal{T}_N^\ell}{argmax} \sum_{j=1}^m \|P_V f_j\|^2, \tag{16}$$

and using the definition of $\mathcal{T}_N^\ell$, we have that

$$\max_{V \in \mathcal{T}_N^\ell} \sum_{j=1}^m \|P_V f_j\|^2 = \max_{\Omega \in M_N^\ell} \sum_{j=1}^m \|P_{V_\Omega} f_j\|^2. \tag{17}$$

So, we need to find $\Omega \in M_N^\ell$ that yields the maximum in (17).

Using Lemma 2.5 we see that for each $\Omega \in M_N^\ell$,

$$\begin{aligned}
\sum_{j=1}^m \|P_{V_\Omega} f_j\|^2 &= \sum_{j=1}^m \|P_{\widehat{V_\Omega}} \widehat{f_j}\|^2 \\
&= \sum_{j=1}^m \int_{\mathcal{U}} \|P_{J_{V_\Omega}}(\omega)(\tau f_j(\omega))\|_{\ell^2(\mathbb{Z}^d)}^2 \, d\omega \\
&= \int_{\mathcal{U}} \sum_{j=1}^m \|P_{J_{V_\Omega}}(\omega)(\tau f_j(\omega))\|_{\ell^2(\mathbb{Z}^d)}^2 \, d\omega.
\end{aligned} \tag{18}$$

Recall that $P_{J_{V_\Omega}}(\omega)$ denotes the orthogonal projection onto the closed subspace $J_{V_\Omega}(\omega)$ of $\ell^2(\mathbb{Z}^d)$.

Furthermore, if $\Omega \in M_N^\ell$, we know from the proof of Proposition 4.3 that $\dim(J_{V_\Omega}(\omega)) = \ell$ for a.e. $\omega \in \mathcal{U}$. Note that $J_{V_\Omega}(\omega)$ agrees with the subspace of $\ell^2(\mathbb{Z}^d)$ of the sequences supported in O_ω. Then there exists a unique set of ℓ integer vectors $\mathbf{k}^\Omega(\omega) = \{k_1^\Omega(\omega), \ldots, k_\ell^\Omega(\omega)\} \subset \mathbb{Z}^d$ such that $\text{span}\{\delta_{k_j^\Omega(\omega)} : j = 1, \ldots, \ell\} = J_{V_\Omega}(\omega)$, for a.e. $\omega \in \mathcal{U}$. Here δ_j denotes the canonical vector in $\ell^2(\mathbb{Z}^d)$. i.e. $\delta_j(s) = 0$ if $s \neq j$ and 1 otherwise. Note that, since $\Omega \subset C_N$ necessarily $\|k_j^\Omega(\omega)\|_\infty \leq N$, for each j and ω. Combining this observation with (18) we obtain,

$$\sum_{j=1}^m \|P_{V_\Omega} f_j\|^2 = \int_{\mathcal{U}} \sum_{j=1}^m \sum_{s=1}^\ell |\widehat{f_j}(\omega + k_s^\Omega(\omega))|^2 \, d\omega. \tag{19}$$

So, now we need to maximize the left hand side in (19) over all the sets $\Omega \in M_N^\ell$.

Note that given $\Omega \in M_N^\ell$, for almost each $\omega \in \mathcal{U}$, the set Ω contains exactly ℓ elements from the sequence $\{\omega + k, k \in \mathbb{Z}^d\}$. Then we can pick for each $\omega \in \mathcal{U}$ (up to a set of zero measure) ℓ translations $k_s^*(\omega)$ such that $\sum_{j=1}^m \sum_{s=1}^\ell |\widehat{f_j}(\omega + k_s^*(\omega))|^2$ is maximum over all sets of ℓ translations $\mathbf{k} = \{k_1, \ldots, k_\ell\} \subset \mathbb{Z}^d$, with $\|k_j\|_\infty \leq N$. The maximum exists since the fibers of f_j are $\ell^2(\mathbb{Z}^d)$-sequences and the number of translations considered is finite.

Call $\mathcal{K}$ the set of admissible translations i.e. $\mathcal{K} = \{\mathbf{k} = \{k_1, \ldots, k_\ell\} \subset \mathbb{Z}^d : \|k_j\|_\infty \leq N\}$ and for $\mathbf{k} \in \mathcal{K}$ set $H_{\mathbf{k}}(\omega) = \sum_{j=1}^m \sum_{s=1}^\ell |\widehat{f_j}(\omega + k_s(\omega))|^2$.

Our goal is to construct a set Ω such that the associated space V_Ω is optimal. So the idea is to construct the optimal set Ω^* considering for each $\omega \in \mathcal{U}$ the optimal translations $\{\omega + k_s^*(\omega) : s = 1, \ldots, \ell\}$, and then taking the union over almost all $\omega \in \mathcal{U}$.

For this we define for each $\mathbf{k} = \{k_1, \ldots, k_\ell\} \in \mathcal{K}$ the following subset of $\mathcal{U}$,

$$E_{\mathbf{k}} = \{\omega \in \mathcal{U} : H_{\mathbf{k}}(\omega) \geq H_{\mathbf{r}}(\omega), \ \forall \ \mathbf{r} = \{r_1, \ldots, r_\ell\} \in \mathcal{K}\},$$

i.e., $E_{\mathbf{k}}$ is the set of $\omega \in \mathcal{U}$ for which the maximum is attained for $\mathbf{k} = \{k_1, \ldots, k_\ell\}$. Note that $E_{\mathbf{k}}$ could be the empty set for some $\mathbf{k} = \{k_1, \ldots, k_\ell\}$ and the sets $E_{\mathbf{k}}$ may not be disjoint.

Finally we define our optimal set as,

$$\Omega^* = \bigcup_{\mathbf{k} \in \mathcal{K}} \bigcup_{j=1}^\ell E_{\mathbf{k}} + k_j.$$

We will now prove that Ω^* is measurable. First we note that $E_{\mathbf{k}}$ is a measurable set for each $\mathbf{k} \in \mathcal{K}$ since,

$$E_{\mathbf{k}} = \bigcap_{\mathbf{r} \in \mathcal{K}} F_{\mathbf{r}}^{\mathbf{k}},$$

where,

$$F_{\mathbf{r}}^{\mathbf{k}} = \{\omega \in \mathcal{U} : H_{\mathbf{k}}(\omega) \geq H_{\mathbf{r}}(\omega)\}.$$

Now, since $F_{\mathbf{r}}^{\mathbf{k}}$ is measurable for all $\mathbf{r} \in \mathcal{K}$, we obtain that $E_{\mathbf{k}}$ is measurable and so is Ω^*.

Furthermore, by construction, Ω^* is in M_N^ℓ. Since for all $\Omega \in M_N^\ell$ we have that,

$$\sum_{j=1}^{m} \sum_{s=1}^{\ell} |\widehat{f}_j(\omega + k_s^\Omega(\omega))|^2 \leq \sum_{j=1}^{m} \sum_{s=1}^{\ell} |\widehat{f}_j(\omega + k_s^*(\omega))|^2 \quad \text{for almost all } \omega \in \mathcal{U},$$

taking the integral over $\mathcal{U}$ we get

$$\sum_{j=1}^{m} \|P_{V_\Omega} f_j\|^2 = \int_{\mathcal{U}} \sum_{j=1}^{m} \sum_{s=1}^{\ell} |\widehat{f}_j(\omega + k_s^\Omega(\omega))|^2 \, d\omega$$

$$\leq \int_{\mathcal{U}} \sum_{j=1}^{m} \sum_{s=1}^{\ell} |\widehat{f}_j(\omega + k_s^*(\omega))|^2 \, d\omega = \sum_{j=1}^{m} \|P_{V_{\Omega^*}} f_j\|^2.$$

This shows that $\Omega^* \in M_N^\ell$ is optimal over all $\Omega \in M_N^\ell$. We conclude that $V_{\Omega^*} \in \mathcal{T}_N^\ell$ is a solution for the data $\mathcal{F}$. $\quad \square$

Remark 4.5. Notice that if Ω_N^* is the optimal multi-tile set for the class $\mathcal{T}_N^\ell$ for some data $\mathcal{F}$, then the approximation error is given by

$$\mathcal{E}_N(\mathcal{F}, \ell) = \int_{\mathbb{R}^d \setminus \Omega_N^*} \sum_{j=1}^{m} |\widehat{f}_j(\omega)|^2 d\omega = \int_{C_N \setminus \Omega_N^*} \sum_{j=1}^{m} |\widehat{f}_j(\omega)|^2 d\omega + \int_{\mathbb{R}^d \setminus C_N} \sum_{j=1}^{m} |\widehat{f}_j(\omega)|^2 d\omega.$$

Clearly $\mathcal{E}_N(\mathcal{F}, \ell) \geq \mathcal{E}_{N+1}(\mathcal{F}, \ell)$. So $\mathcal{E}(\mathcal{F}, \ell) := \lim_{N \to \infty} \mathcal{E}_N(\mathcal{F}, \ell)$ is somehow the optimal error. Since $\mathcal{F} \subset L^2(\mathbb{R}^d)$ then the second integral goes to zero when N goes to infinite, for functions with good decay at infinite we will be close to the optimal error for conveniently large N.

Remark 4.6. In Proposition 4.3 we show, for an element of $\mathcal{T}^\ell$, how to construct a set of generators that gives a Riesz basis of translates in $\mathbb{Z}^d$. There are many ways to construct other sets of generators that gives Riesz basis of translates. Recently Grepstad–Lev in [14] constructed a basis of exponentials for $L^2(\Omega)$ when $\Omega \subset \mathbb{R}^d$ is a multi-tile. Later on, Kolountzakis [19] gave a simpler proof of this result in a slightly more general form. Precisely they prove the following result.

Theorem 4.7. *(See Theorem 1 of [19].) Suppose $\Omega \subset \mathbb{R}^d$ is bounded, measurable and multi-tiles $\mathbb{R}^d$ when translated by $\mathbb{Z}^d$ at level ℓ. Then there exist vectors $a_1, \ldots, a_\ell \in \mathbb{R}^d$ such that the exponentials*

$$e^{-2\pi i (a_j + k)\omega} \quad j = 1, \ldots, \ell, \quad k \in \mathbb{Z}^d$$

form a Riesz basis for $L^2(\Omega)$.

From Theorem 4.7, we can obtain immediately a set of generators for V_Ω. Let φ be such that $\widehat{\varphi} = \chi_\Omega$. If $a_1, \ldots, a_\ell \in \mathbb{R}^d$ are as in Theorem 4.7, then $V_\Omega = S(\varphi_1, \ldots, \varphi_\ell)$ with $\varphi_j = t_{a_j}\varphi, \ j = 1, \ldots, \ell$, and the translates of $\varphi_1, \ldots, \varphi_\ell$ form a Riesz basis of V_Ω.

In general, all the Riesz basis for V_Ω can be described in the following way:

Let $A = \{a_{js}\} \in [L^2(\mathcal{U})]^{\ell \times \ell}$ be a measurable matrix, such that $0 < c_1 \leq \lambda(\omega) \leq c_2$ for every eigenvalue $\lambda(\omega)$ and for almost each $\omega \in \mathcal{U}$. Set $k(\omega) = (k_1(\omega), \ldots, k_\ell(\omega))$ such that $w + k_s(\omega) \in \Omega$. Define φ_j such that $\widehat{\varphi}_j(\omega + k_s(w)) = a_{js}(\omega)$. Using the results stated in subsection 2.1 it is not difficult to see that $\varphi_1, \ldots, \varphi_\ell$ are measurable, $V_\Omega = S(\varphi_1, \ldots, \varphi_\ell)$ and the translates of $\varphi_1, \ldots, \varphi_\ell$ form a Riesz basis of V_Ω.

5. The discrete case

The optimal subspace V^* in Theorem 3.1 is the closest to the data $\mathcal{F}$ over all subspaces V in the class $\mathcal{V}_M^\ell$. It is not difficult to see that almost each fiber space $J_{V^*}(\omega) \subset \ell^2(\mathbb{Z}^d)$ of V^* is the closest to the fibers of our data, $\tau(\mathcal{F})(\omega) = \{\tau f_1(\omega), \ldots, \tau f_m(\omega)\}$ over a certain class of closed subspaces of $\ell^2(\mathbb{Z}^d)$ that we will call $\mathcal{D}_\mathcal{N}^\ell$. Clearly this class of subspaces is determine by the class $\mathcal{V}_M^\ell$. So, Theorem 3.1, implies an approximation result in $\ell^2(\mathbb{Z}^d)$, for a very particular class determined by the extra-invariance. Therefore it is interesting to see if this approximation result in $\ell^2(\mathbb{Z}^d)$, extends to more general classes. We will define in what follows a very general class $\mathcal{D}_\mathcal{N}^\ell$. The cases coming from the continuous case will be particular cases of our general definition. The proof of the optimality that we obtain is more general and cannot follow from Theorem 3.1.

In what follows, the Hilbert space we consider is $\ell^2(\mathbb{Z}^d)$. We define the class of approximating subspaces in the following way:

Let $\mathcal{N}$ be an arbitrary finite set and $\{D_\sigma : \sigma \in \mathcal{N}\}$ a partition of $\mathbb{Z}^d$, that is, $\mathbb{Z}^d = \bigcup_{\sigma \in \mathcal{N}} D_\sigma$, where the union is disjoint.

For $a \in \ell^2(\mathbb{Z}^d)$, we denote $a^\sigma = \mathbf{1}_{\mathbf{D}_\sigma} a$, where $\mathbf{1}_{\mathbf{D}_\sigma}$ denotes the indicator of D_σ. Given $S \subset \ell^2(\mathbb{Z}^d)$ a closed subspace we define

$$S_\sigma = \{a^\sigma : a \in S\}, \quad \text{for each } \sigma \in \mathcal{N}.$$

We define the class of approximating subspaces by

$$\mathcal{D}_\mathcal{N}^\ell = \{S \subset \ell^2(\mathbb{Z}^d) : S \text{ is a subspace, } \dim(S) \leq \ell \text{ and } S_\sigma \subset S, \ \forall \sigma \in \mathcal{N}\}. \tag{20}$$

Note that $S \in \mathcal{D}_\mathcal{N}^\ell$ if and only if $\dim(S) \leq \ell$ and S is the orthogonal sum of the subspaces S_σ i.e.,

$$S = \oplus_{\sigma \in \mathcal{N}} S_\sigma.$$

For a given set $\mathcal{A} = \{a_1, \ldots, a_m\} \subset \ell^2(\mathbb{Z}^d)$ consider for each, $\sigma \in \mathcal{N}$, the Gramian matrix $G_\sigma \in \mathbb{C}^{m \times m}$ of the data $\mathcal{A}_\sigma = \{a_1^\sigma, \ldots, a_m^\sigma\}$, that is $(G_\sigma)_{k,l} = \langle a_k^\sigma, a_l^\sigma, \rangle, \ k, l = 1, \ldots, m$, with eigenvalues $\lambda_1^\sigma \geq \cdots \geq \lambda_m^\sigma$, and orthonormal corresponding left eigenvectors $y_1^\sigma, \ldots, y_m^\sigma$.

Now set $\Lambda = \{\lambda_j^\sigma : j = 1, \cdots m, \ \sigma \in \mathcal{N}\}$ and collect in Λ_ℓ the ℓ first biggest eigenvalues of Λ that is if $\lambda \in \Lambda_\ell$ then $\lambda \geq \mu$ for all $\mu \in \Lambda \setminus \Lambda_\ell$.

Write $\Lambda_\ell = \{\lambda_1, \ldots, \lambda_\ell\}$. For each $s = 1, \ldots, \ell$ we define the sequence $q_s \in \ell^2(\mathbb{Z}^d)$ in the following way:

Since $\lambda_s = \lambda_{j_s}^{\sigma_s}$ for some $\sigma_s \in \mathcal{N}$ and some $j_s = 1, \ldots, m$, then λ_s is an eigenvalue of G_{σ_s}. Let $y_{j_s}^{\sigma_s}$ be the corresponding left eigenvector $y_{j_s}^{\sigma_s} = (y_{j_s}^{\sigma_s}(1), \ldots, y_{j_s}^{\sigma_s}(m))$.

Then define, if $\lambda_s \in \Lambda_\ell, \lambda_s \neq 0$

$$q_s := (\lambda_s)^{-1/2} \sum_{k=1}^m y_{j_s}^{\sigma_s}(k) a_k^{\sigma_s}. \tag{21}$$

If $\lambda_s = 0$ we define q_s to be the zero sequence.

With this notation we can state the main theorem of this section:

Theorem 5.1. *Let $m, \ell \in \mathbb{N}$ and $\mathcal{N}$ a finite set. Assume that a set $\mathcal{A} = \{a_1, \ldots, a_m\} \subset \ell^2(\mathbb{Z}^d)$ is given. Then there exists $S^* \in \mathcal{D}_{\mathcal{N}}^{\ell}$ that satisfies*

$$\sum_{j=1}^{m} \|a_j - P_{S^*} a_j\|^2 \leq \sum_{j=1}^{m} \|a_j - P_S a_j\|^2, \quad \forall S \in \mathcal{D}_{\mathcal{N}}^{\ell}. \tag{22}$$

Moreover, we have that

(1) *$S^* = span\{q_1, \ldots, q_\ell\}$ where $q_1, \ldots, q_\ell$ are defined in (21). Also, the vectors $\{q_1, \ldots, q_\ell\}$ form a Parseval frame for S^*.*
(2) *The error in the approximation is*

$$\mathcal{E}(\mathcal{A}, \mathcal{N}, \ell) = \sum_{\lambda \in \Lambda \setminus \Lambda_\ell} \lambda.$$

Proof. First, we observe that (22) is equivalently to,

$$\sum_{j=1}^{m} \|P_{S^*} a_j\|^2 \geq \sum_{j=1}^{m} \|P_S a_j\|^2, \quad \forall S \in \mathcal{D}_{\mathcal{N}}^{\ell}.$$

Furthermore, if $S \in \mathcal{D}_{\mathcal{N}}^{\ell}$ then

$$\sum_{j=1}^{m} \|P_S a_j\|^2 = \sum_{j=1}^{m} \| \sum_{\sigma \in \mathcal{N}} P_{S_\sigma} a_j\|^2 = \sum_{j=1}^{m} \| \sum_{\sigma \in \mathcal{N}} P_{S_\sigma} a_j^\sigma\|^2 = \sum_{\sigma \in \mathcal{N}} \sum_{j=1}^{m} \|P_{S_\sigma} a_j^\sigma\|^2,$$

where $a_j = \sum_{\sigma \in \mathcal{N}} a_j^\sigma$.

In order to construct an optimal subspace $S^* \in \mathcal{D}_{\mathcal{N}}^{\ell}$ for the data $\mathcal{A}$, we will find an optimal subspace S_σ for each $\sigma \in \mathcal{N}$, of dimension at most $\alpha_\sigma \in \{1, \ldots, \ell\}$ for the data $\mathcal{A}_\sigma = \{a_1^\sigma, \ldots, a_m^\sigma\}$. The existence of the optimal subspaces are provided by Theorem 4.1 of [2]. We need $\sum_{\sigma \in \mathcal{N}} \dim(S_\sigma) = \sum_{\sigma \in \mathcal{N}} \alpha_\sigma = \dim(S^*) \leq \ell$.

Thus if $\mathcal{Q} = \{\alpha = \{\alpha_\sigma\} : 0 \leq \alpha_\sigma \leq \ell$ and $\sum_{\sigma \in \mathcal{N}} \alpha_\sigma \leq \ell\}$, then for each choice of $\alpha \in \mathcal{Q}$ we will find optimal subspaces $\{S_\sigma^\alpha : \sigma \in \mathcal{N}\}$ and define $S^\alpha = \oplus_{\sigma \in \mathcal{N}} S_\sigma^\alpha$.

The candidate for S^* is the space S^α which minimize the expression (22) over all $\alpha \in \mathcal{Q}$. Let $\beta \in \mathcal{Q}$ be the minimizer. Hence β satisfies,

$$\sum_{\sigma \in \mathcal{N}} \sum_{j=1}^{m} \|P_{S_\sigma^\alpha} a_j^\sigma\|^2 \leq \sum_{\sigma \in \mathcal{N}} \sum_{j=1}^{m} \|P_{S_\sigma^\beta} a_j^\sigma\|^2 \quad \forall \, \alpha \in \mathcal{Q}. \tag{23}$$

Therefore, the subspace $S^* := S^\beta = \oplus_{\sigma \in \mathcal{N}} S_\sigma^\beta$ is the optimal subspace we need. It is straightforward to see that $S^* \in \mathcal{D}_{\mathcal{N}}^{\ell}$ and that S^* is optimal.

Using Theorem 4.1 of [2] we obtain that for each $\alpha \in \mathcal{Q}$, the error of approximation for the data $\mathcal{A}_\sigma$ and the class of subspaces of dimension at most α_σ is given by

$$\mathcal{E}(\mathcal{A}_\sigma, \alpha_\sigma) = \sum_{s=\alpha_\sigma+1}^{m} \lambda_s^\sigma.$$

So the distance between the α-optimal subspace S^α and the data $\mathcal{A}$ is,

$$E(\alpha) = \sum_{\sigma \in \mathcal{N}} \mathcal{E}(\mathcal{A}_\sigma, \alpha_\sigma) = \sum_{\sigma \in \mathcal{N}} \sum_{s=\alpha_\sigma+1}^{m} \lambda_s^\sigma. \tag{24}$$

Let κ be the number of elements in $\mathcal{N}$. We see that $E(\alpha)$ is minimum when the $m\kappa - \ell$ eigenvalues used in (24) are the smallest from the set $\Lambda = \{\lambda_j^\sigma : j = 1, \ldots, m, \ \sigma \in \mathcal{N}\}$. Therefore if we set $\Lambda_\ell \subset \Lambda$ the set of the ℓ biggest eigenvalues from Λ, the optimal $\beta = \{\beta_\sigma\} \in \mathcal{Q}$ satisfies that

$$\bigcup_{\sigma \in \mathcal{N}} \{\lambda_1^\sigma, \ldots, \lambda_{\beta_\sigma}^\sigma\} = \Lambda_\ell.$$

Therefore,

$$\mathcal{E}(\mathcal{A}, \mathcal{N}, \ell) = \sum_{\sigma \in \mathcal{N}} \mathcal{E}(\mathcal{A}_\sigma, \beta_\sigma) = \sum_{\sigma \in \mathcal{N}} \sum_{j=1}^{m} \|a_j^\sigma - P_{S_\sigma^\beta} a_j^\sigma\|^2 = \sum_{\sigma \in \mathcal{N}} \sum_{s=\beta_\sigma+1}^{m} \lambda_s^\sigma = \sum_{\lambda \in \Lambda \setminus \Lambda_\ell} \lambda.$$

In order to construct the generators of S^β it is enough to construct the generators of each S_σ^β. Since S_σ^β are optimal subspaces for the data $\mathcal{A}_\sigma$ according with Theorem 4.1 of [2] the generators of S_σ^β are given by (21). That is the set $\{q_s : \sigma_s = \sigma\}$ is a Parseval frame of S_σ^β.

Since the subspaces S_σ^β are mutually orthogonal, $\{q_1, \ldots, q_\ell\}$ is a set of Parseval frame generators for the optimal space $S^* = S^\beta$. $\quad\square$

Remark 5.2. As explained at the beginning of this section there is a reason to consider this particular class of subspaces for the discrete case. If a SIS V is M extra-invariant for some proper subgroup M of $\mathbb{R}^d$ containing $\mathbb{Z}^d$, then its fiber spaces $J_V(\omega)$ satisfies exactly the conditions that we imposed on the class $\mathcal{D}_\mathcal{N}^\ell$ where the partition of $\mathbb{Z}^d$ is $\{B_\sigma \cap \mathbb{Z}^d : \sigma \in \mathcal{N}\}$ and B_σ and $\mathcal{N}$ are as in (5). So, the discrete result (Theorem 5.1) provides a different proof of Theorem 3.1 using properties of range functions (see Definition 2.3), without the need of Theorem 2.7. Actually this proof includes Theorem 2.7.

Acknowledgments

References

[1] A. Aldroubi, M. Anastasio, C. Cabrelli, U. Molter, A dimension reduction scheme for the computation of optimal unions of subspaces, Sampl. Theory Signal Image Process. 10 (1–2) (2011) 135–150.
[2] A. Aldroubi, C. Cabrelli, D. Hardin, U. Molter, Optimal shift invariant spaces and their Parseval frame generators, Appl. Comput. Harmon. Anal. 23 (2) (2007) 273–283.
[3] A. Aldroubi, C. Cabrelli, C. Heil, K. Kornelson, U. Molter, Invariance of a shift-invariant space, J. Fourier Anal. Appl. 16 (1) (2010) 60–75.
[4] A. Aldroubi, C. Cabrelli, U. Molter, Optimal non-linear models for sparsity and sampling, J. Fourier Anal. Appl. 14 (5–6) (2008) 793–812.
[5] A. Aldroubi, K. Gröchenig, Nonuniform sampling and reconstruction in shift-invariant spaces, SIAM Rev. 43 (4) (2001) 585–620 (electronic).
[6] A. Aldroubi, I. Krishtal, R. Tessera, H. Wang, Principal shift-invariant spaces with extra invariance nearest to observed data, Collect. Math. 63 (3) (2012) 393–401.

[7] A. Aldroubi, R. Tessera, On the existence of optimal unions of subspaces for data modeling and clustering, Found. Comput. Math. 11 (3) (2011) 363–379.

[8] M. Anastasio, C. Cabrelli, V. Paternostro, Invariance of a shift-invariant space in several variables, Complex Anal. Oper. Theory 5 (4) (2011) 1031–1050.

[9] M. Bownik, The structure of shift-invariant subspaces of $L^2(\mathbb{R}^n)$, J. Funct. Anal. 177 (2) (2000) 282–309.

[10] C. Cabrelli, V. Paternostro, Shift-invariant spaces on LCA groups, J. Funct. Anal. 258 (6) (2010) 2034–2059.

[11] O. Christensen, An introduction to frames and Riesz bases, in: Applied and Numerical Harmonic Analysis, Birkhäuser Boston Inc., Boston, MA, 2003.

[12] C. de Boor, R.A. DeVore, A. Ron, Approximation from shift-invariant subspaces of $L_2(\mathbb{R}^d)$, Trans. Amer. Math. Soc. 341 (2) (1994) 787–806.

[13] C. de Boor, R.A. DeVore, A. Ron, The structure of finitely generated shift-invariant spaces in $L_2(\mathbb{R}^d)$, J. Funct. Anal. 119 (1) (1994) 37–78.

[14] S. Grepstad, N. Lev, Multi-tiling and Riesz bases, Adv. Math. 252 (2014) 1–6.

[15] K. Gröchenig, Foundations of time-frequency analysis, in: Applied and Numerical Harmonic Analysis, Birkhäuser Boston Inc., Boston, MA, 2001.

[16] C. Heil, A basis theory primer, in: Applied and Numerical Harmonic Analysis, expanded edition, Birkhäuser/Springer, New York, 2011.

[17] H. Helson, Lectures on Invariant Subspaces, Academic Press, New York, 1964.

[18] E. Hernández, G. Weiss, A First Course on Wavelets, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1996, with a foreword by Yves Meyer.

[19] M. Kolountzakis, Multiple lattice tiles and Riesz bases of exponentials, Proc. Amer. Math. Soc. 143 (2) (2015) 741–747.

[20] S.G. Mallat, Multiresolution approximations and wavelet orthonormal bases of $L_2(\mathbb{R})$, Trans. Amer. Math. Soc. 315 (1) (1989) 69–87.

[21] A. Ron, Z. Shen, Frames and stable bases for shift-invariant subspaces of $L_2(\mathbb{R}^d)$, Canad. J. Math. 47 (5) (1995) 1051–1094.

[22] R. Tessera, H. Wang, Uncertainty principles in finitely generated shift-invariant spaces with additional invariance, J. Math. Anal. Appl. 410 (1) (2014) 134–143.