Noname manuscript No.
(will be inserted by the editor)

Sparse Matrices in Frame Theory

Felix Krahmer · **Gitta Kutyniok** · **Jakob Lemvig**

Received: date / Accepted: date

Abstract Frame theory is closely intertwined with signal processing by providing a canon of methodologies for the analysis of signals using (redundant) linear measurements. The dual frame associated with a frame then provides a means for reconstruction by a least squares approach. The novel paradigm of sparsity entered this area lately in various ways. Of those, in this survey paper, we will focus on the frames and dual frames which can be written as sparse matrices. The objective for this approach is to ensure not only low-complexity computations, but also high compressibility. We will discuss both existence results as well as explicit constructions.

Keywords Dual Frames · Frames · Redundancy · Signal Processing · Sparse Matrices · Tight Frames

F. Krahmer
Universität Göttingen
Institut für Numerische und Angewandte Mathematik
37083 Göttingen, Germany
Tel.: +49-551-3910584
Fax: +49-551-393944
E-mail: f.krahmer@math.uni-goettingen.de

G. Kutyniok
Technische Universität Berlin
Institut für Mathematik
10623 Berlin, Germany
Tel.: +49-30-314-25758
Fax: +49-30-314-21604
E-mail: kutyniok@math.tu-berlin.de

J. Lemvig
Technical University of Denmark
Department of Mathematics
2800 Kgs. Lyngby, Denmark
Tel.: +45-45253065
Fax: +45-45881399
E-mail: j.lemvig@mat.dtu.dk

1 Introduction

Efficient signal analysis and processing is one of the great scientific challenges to date due to the ever increasing enormous amounts of data human kind is generating. Recently, sparsity has entered the stage as a novel paradigm allowing the development of highly efficient methodologies under certain typically satisfiable constraints. The main objectives of this approach are dimension reduction allowing precise reconstruction of the original signal, high compressibility of data, low-complexity computations, and many more.

1.1 Mathematical Signal Processing

In mathematical signal processing, one canonical first step is the computation of linear (adaptive or non-adaptive) measurements associated with a carefully designed representation system. To be precise, for a signal $x \in \mathbb{R}^n$ and a spanning sequence $(\varphi_i)_{i=1}^m \subset \mathbb{R}^n$, we compute coefficients $c \in \mathbb{R}^m$ by

$$x \mapsto c := \Phi^* x = (\langle x, \varphi_i \rangle)_{i=1}^m \qquad \text{with } \Phi = (\varphi_1 | \ldots | \varphi_m) \in \mathbb{R}^{n \times m}. \tag{1.1}$$

The case when $m = n$ and $(\varphi_i)_{i=1}^m$ constitutes an orthonormal basis for $\mathbb{R}^n$, e.g., when $(\varphi_i)_{i=1}^m$ is simply the Dirac basis, is very well studied. More intriguing are the following two fundamentally different, more recently considered objectives for the transform in (1.1), which go beyond the setting of orthonormal bases. One objective is *dimension reduction*, i.e., $m < n$, of which compressed sensing is a prominent representative. Another objective is *robust analysis*, i.e., $m > n$, which frame theory focuses on.

Both cases face the crucial question of whether it is possible to recover x from the measurement coefficients $c = \Phi^* x$. In the undercomplete case ($m < n$), often convex optimization is used, such as for instance in compressed sensing (see for example the survey paper [16] and the book [12]). In the overcomplete case ($m > n$), least squares is a typical approach, which also frame theory traditionally follows. Thus reconstruction is typically performed by computing

$$((\Phi\Phi^*)^{-1}\Phi)c. \tag{1.2}$$

In this survey paper, we from now on focus on the overcomplete scenario with the objective of deriving a robust analysis. However, with sparsity entering the picture, as we will see, least squares is not always the preferred method of reconstruction.

1.2 Frame Theory

Frame theory – the theory of overcomplete (redundant) Bessel systems – dates back to work by Duffin and Schaeffer in 1952 on non-harmonic Fourier series (see [11]). Its success story in signal processing started in the 90th with the seminal work by Daubechies, Grossman and Meyer [9]. At that time, it was recognized that not only does redundancy of $(\varphi_i)_{i=1}^m$ ensure robustness of $(\langle x, \varphi_i \rangle)_{i=1}^m$ against noise or erasures, but the restrictions of forming an orthonormal basis are often too strong for the construction of many systems. Since the groundbreaking work [9] was published, frames have become a key methodology in signal processing. It should be emphasized that frame theory is not only widely used in the finite dimensional setting (cf. [8]), but also in infinite dimensions. In this survey paper we though focus on finite frames.

In signal processing with frames, three main steps can be identified. The first step is *encoding* or *decomposition* with frames, which is performed by the mapping in (1.1). Main objectives for this step are the design of frames Φ that can be easily stored and allow a low-complexity computation of $\Phi^* x$.

The second step is the *analysis* of the signal based on the frame coefficients $c = \Phi^* x$. Depending on the processing goal (feature detection, inpainting, transmission etc.), the frame Φ needs to be designed accordingly, for instance, by encoding the sought features in the large coefficients. Another main issue in signal processing are linear or non-linear operators $A : \mathbb{R}^m \to \mathbb{R}^m$ applied to the measurements $\Phi^* x$ during the analysis or transmission process. Examples for such operators are erasure operators associated with a diagonal matrix with diagonal entries either 0 or 1, thresholding operators which set to zero all entries smaller (in absolute value) than a given value, or the operator which takes the absolute value of each entry leading to the problem of phase retrieval. Hence one goal then is to minimize the error $\max_{\|x\|_2=1} \|x - ((\Phi\Phi^*)^{-1}\Phi) A \Phi^* x\|_2$.

The third step consists in *reconstruction* of the original or manipulated signal, which is for instance, done by (1.2), where c could be either the original or modified frame coefficients depending on the analysis step. Again, low-complexity computations are one concern. Also, it is often desirable to choose a frame Ψ, a so-called *dual frame*, different from $(\Phi\Phi^*)^{-1}\Phi$ for reconstruction. The reconstruction is then performed by computing Ψc, and one can imagine that again design questions need to be faced.

This shows the richness of the tasks in frame theory, and it is easily imaginable that this richness reflects on the number of commonly used frames available to date. Not wanting to delve too much into details, we just mention equiangular frames, harmonic frames, Gabor frames, wavelet frames, or shearlet frames.

1.3 Desideratum: Sparsity

Sparsity has become an important paradigm in both numerical linear algebra and signal processing. The sparsity of a vector $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ is measured by

$$\|x\|_0 := \#\{i \in \{1, \dots, n\} : x_i \neq 0\},$$

and x is called *k-sparse*, if $\|x\|_0 \leq k$. Along the same lines, sparsity of a matrix means that many of the matrix entries vanish, i.e., the quantity

$$\|\Phi\|_0 := \#\{(i,j) : \varphi_{i,j} \neq 0\}, \qquad \text{for } \Phi = [\varphi_{i,j}] \in \mathbb{R}^{n \times m}. \tag{1.3}$$

being small. Sparsity nowadays plays two conceptually very different roles. On the one hand, sparse representations guarantee efficient storage and processing of data.

For example, multiplying a vector with a sparse matrix requires less operations and the sparse representation of a signal can be directly used for efficient compression. On the other hand, sparse representations epitomize structural simplicity. In fact, many problems in signal processing are intrinsically ill-posed, so only such structural assumptions on the solution make accurate and stable solutions possible. Key examples are in the theory of compressed sensing as introduced in parallel by Donoho [10] and Candès, Romberg, and Tao [4] in 2006. In this context, also weaker versions of sparsity such as compressibility or the behavior of the error of best n-term approximation are frequently used.

Both these utilization paradigms of the concept of sparsity have direct connections to frame theory. For the first one, this connection can be drawn in two different ways. Firstly, one can aim to design frames and dual frames to form sparse matrices themselves, with the goal of an efficient encoding and reconstruction process. On the other hand, the additional freedom resulting from the redundancy when working with frame representations rather than basis representations can allow to derive sparser representations of a vector. The second one appears in frame theory as a design question for the measurement matrix in compressed sensing, which computes Φx for some x in the higher dimensional space $\mathbb{R}^m$.

In this survey, we focus on the very first connection, that is, sparsity of the frame and dual frame matrices as a means to ensure a more efficient encoding and reconstruction process. However, the inverse problems viewpoint on sparsity also plays an important role, as finding sparse frames under certain constraints has many structural similarities to finding sparse solutions to inverse problems.

Frames as sparse matrices were first analyzed in 2011 in [7], whereas sparse duals were first considered and discussed in [14]. For both situations, results on the optimal sparsity were derived, existence results were proven, and algorithms provided.

1.4 Outline

In this paper, we first provide an introduction to frame theory (Section 2). In Section 3, we then focus on sparse frames, and discuss optimality results, existence results as well as explicit constructions. Similar considerations will be undertaken in Section 4, then focusing on associated dual frames.

2 Basics of Frame Theory

We start with reviewing the basic definitions and notations of frame theory, which will be used in the sequel.

2.1 Frames

A frame is a family of vectors which ensures stability of the map introduced in (1.1). The precise definition is as follows. We also emphasize that for simplicity, we present the definition as well as the results solely for the real case. Notice that most results also hold for the complex case.

Definition 2.1 A family of vectors $(\varphi_i)_{i=1}^m$ in $\mathbb{R}^n$ is called a *frame for* $\mathbb{R}^n$, if there exist constants $0 < A \leq B < \infty$ such that

$$A\|x\|^2 \leq \sum_{i=1}^{M} |\langle x, \varphi_i \rangle|^2 \leq B\|x\|^2 \quad \text{for all } x \in \mathbb{R}^n. \tag{2.1}$$

The constants A and B are called *lower and upper frame bound* for the frame, respectively. If $A = B$ is possible in (2.1), then $(\varphi_i)_{i=1}^M$ is called an *A-tight frame*, and if $A = B = 1$, a *Parseval frame*. Note that in the latter case, the system satisfies the Parseval equality. If there exists a constant c such that $\|\varphi_i\| = c$ for all $i = 1, 2, \ldots, m$, then $(\varphi_i)_{i=1}^m$ is an *equal-norm frame*. If $c = 1$, $(\varphi_i)_{i=1}^M$ is a *unit-norm frame*. Finally, the values $(\langle x, \varphi_i \rangle)_{i=1}^m$ are called the *frame coefficients* of the vector x with respect to the frame $(\varphi_i)_{i=1}^m$.

For a given frame $\Phi = (\varphi_i)_{i=1}^m$ and a fixed orthonormal basis $(e_j)_{j=1}^n$, we let Φ denote the $n \times m$ *frame matrix*, whose ith column is the coefficient vector of φ_i. Note that with a small abuse of notation, we will denote by Φ both the frame and the corresponding frame matrix. The condition (2.1) for a frame then reads

$$A\|x\|^2 \leq \|\Phi^* x\|_2^2 \leq B\|x\|^2 \quad \text{for all } x \in \mathbb{R}^n.$$

Trivial, but useful observations, are combined in the following result. Since the proofs are straightforward, we leave them to the reader.

Lemma 2.2 *Let $(\varphi_i)_{i=1}^m$ be a family of vectors in $\mathbb{R}^n$.*

(i) If $(\varphi_i)_{i=1}^m$ is an orthonormal basis, then $(\varphi_i)_{i=1}^m$ is a Parseval frame. The converse is not true in general.
(ii) $(\varphi_i)_{i=1}^m$ is a unit-norm Parseval frame if and only if it is an orthonormal basis.

An easy example of a equal-norm Parseval frame, which does not constitute an orthonormal basis, is the frame in $\mathbb{R}^2$ given by

$$\left\{ \sqrt{\tfrac{2}{3}} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \sqrt{\tfrac{2}{3}} \begin{pmatrix} \frac{\sqrt{3}}{2} \\ -\frac{1}{2} \end{pmatrix}, \sqrt{\tfrac{2}{3}} \begin{pmatrix} -\frac{\sqrt{3}}{2} \\ -\frac{1}{2} \end{pmatrix} \right\}$$

Due to the shape of its vectors in $\mathbb{R}^2$, this frame is typically called *Mercedes-Benz frame*.

2.2 The Frame Operator

Given a frame $(\varphi_i)_{i=1}^m$, its signal processing performance is crucially determined by the following three operators. We remark that the first operator was already introduced in (1.1).

Definition 2.3 Let $\Phi = (\varphi_i)_{i=1}^m$ be a frame in $\mathbb{R}^n$.

(i) The associated *analysis operator* $T : \mathbb{R}^n \to \mathbb{R}^m$ is defined by

$$Tx := \Phi^* x = \big(\langle x, \varphi_i \rangle\big)_{i=1}^m, \quad x \in \mathbb{R}^n.$$

(ii) The associated *synthesis operator* is defined to be the adjoint operator T^*. A short computation shows that $T^* : \mathbb{R}^m \to \mathbb{R}^n$ is given by

$$T^* c := \Phi c = \sum_{i=1}^m c_i \varphi_i, \quad c = (c_i)_{i=1}^m \in \mathbb{R}^m.$$

(iii) The associated *frame operator* is the map $S : \mathbb{R}^n \to \mathbb{R}^n$ defined to be

$$Sx := T^* T x = \Phi \Phi^* x = \sum_{i=1}^m \langle x, \varphi_i \rangle, \quad x \in \mathbb{R}^n.$$

The following result can be easily proved by linear algebra techniques.

Theorem 2.4 *The frame operator S of a frame $(\varphi_i)_{i=1}^m$ for $\mathbb{R}^n$ with frame bounds A and B is a positive, self-adjoint invertible operator satisfying*

$$A \cdot I_n \leq S \leq B \cdot I_n.$$

2.3 Reconstruction Strategy

We next aim to reconstruct the original signal from the image under the analysis operator. This can be achieved by application of a linear operator by using Theorem 2.4 as the following result shows.

Theorem 2.5 *Let $\Phi = (\varphi_i)_{i=1}^m$ be a frame for $\mathbb{R}^n$ with frame operator S. Then, for every $x \in \mathbb{R}^n$, we have*

$$x = S^{-1} S x = ((\Phi \Phi^*)^{-1} \Phi)(\Phi^* x) = \sum_{i=1}^m \langle x, \varphi_i \rangle S^{-1} \varphi_i.$$

This is the well-known least squares reconstruction. Due to the redundancy of the frame, also other matrices $\Psi \in \mathbb{R}^{n \times m}$ do exist which satisfy the reconstruction condition $\Psi \Phi^* = I_n$ with I_n being the identity matrix on $\mathbb{R}^n$. For those, the following terminology is common in frame theory.

Definition 2.6 Let $\Phi = (\varphi_i)_{i=1}^m$ be a frame for $\mathbb{R}^n$ with frame operator S. Then the system $(S^{-1} \varphi_i)_{i=1}^m = (\Phi \Phi^*)^{-1} \Phi$ is called the *canonical dual frame*. In general, every frame $\Psi = (\psi_i)_{i=1}^m$ for $\mathbb{R}^n$ satisfying

$$x = \Psi \Phi^* x = \sum_{i=1}^m \langle x, \varphi_i \rangle \psi_i, \quad x \in \mathbb{R}^n$$

is referred to as an *(alternate) dual frame for Φ*.

The set of all dual frames can be explicitly expressed by using classical linear algebra.

Proposition 2.7 *Let $\Phi = (\varphi_i)_{i=1}^m$ be a frame for $\mathbb{R}^n$. Then, every dual frame $\Psi = (\psi_i)_{i=1}^m$ for Φ is of the form*

$$\Psi = (\Phi\Phi^*)^{-1}\Phi + L(I_m - \Phi^*(\Phi\Phi^*)^{-1}\Phi), \quad \text{where } L \in \mathbb{R}^{n\times m}. \tag{2.2}$$

We remark that the canonical dual frame plays no special role in (2.2) in the sense that, if $\widetilde{\Psi}$ is just *some* dual, then all duals are obtained by

$$\Psi = \widetilde{\Psi} + L(I_m - \Phi^*\widetilde{\Psi}),$$

where $L \in \mathbb{R}^{n\times m}$. If Φ forms an A-tight frame, by Theorem 2.4 the frame operator is a multiple of the identity, which then leads to the reconstruction formula

$$x = A^{-1}\Phi(\Phi^*x) = A^{-1}\sum_{i=1}^m \langle x, \varphi_i\rangle \varphi_i.$$

In this case, the canonical frame coincides with $A^{-1}\Phi$. This shows that from a signal processing perspective, tight frame are highly desirable.

2.4 Expansion in Frames

From Theorem 2.4, we can also deduce a different formula, which can be regarded as an expansion of x in terms of the frame $(\varphi_i)_{i=1}^m$.

Theorem 2.8 *Let $\Phi = (\varphi_i)_{i=1}^m$ be a frame for $\mathbb{R}^n$ with frame operator S. Then, for every $x \in \mathbb{R}^n$, we have*

$$x = SS^{-1}x = \Phi((\Phi\Phi^*)^{-1}\Phi^*)x = \sum_{i=1}^m \langle x, S^{-1}\varphi_i\rangle \varphi_i.$$

The specifically chosen sequence of coefficients $(\Phi\Phi^*)^{-1}\Phi^*x$ is the one being minimal in the ℓ_2 norm among all coefficient sequences.

Proposition 2.9 *Let $\Phi = (\varphi_i)_{i=1}^m$ be a frame for $\mathbb{R}^n$ with frame operator S, and let $x \in \mathbb{R}^n$. If $c = (c_i)_{i=1}^m$ are scalars such that $x = \Phi c = \sum_{i=1}^m c_i\varphi_i$, then*

$$\begin{aligned}\|c\|_2^2 &= \|(\Phi\Phi^*)^{-1}\Phi^*x\|_2^2 + \|c - (\Phi\Phi^*)^{-1}\Phi^*x\|_2^2 \\ &= \sum_{i=1}^M |\langle x, S^{-1}\varphi_i\rangle|^2 + \sum_{i=1}^M |c_i - \langle x, S^{-1}\varphi_i\rangle|^2.\end{aligned}$$

Proposition 2.9 immediately implies that for any sequence $c = (c_i)_{i=1}^m$ satisfying $x = \sum_{i=1}^m c_i\varphi_i$, we have

$$\sum_{i=1}^M |\langle x, S^{-1}\varphi_i\rangle|^2 = \|(\Phi\Phi^*)^{-1}\Phi^*x\|_2^2 < \|c\|_2^2$$

unless $c = (\Phi\Phi^*)^{-1}\Phi^*x$.

2.5 Construction of Tight Frames

As already debated, due to their advantageous reconstruction properties, it is particularly desirable to construct tight frames. We distinguish between two different situations: If a given frame shall be modified to become a tight frame, or if only specific parameters are given according to which a tight frame shall be constructed.

If a frame is already given, there exists a very straightforward way to modify it to even become Parseval.

Proposition 2.10 *Let $\Phi = (\varphi_i)_{i=1}^m$ be a frame for $\mathbb{R}^n$ with frame operator S. Then*

$$(S^{-1/2}\varphi_i)_{i=1}^m = (\Phi\Phi^*)^{-1/2}\Phi^*$$

forms a Parseval frame.

Proof This follows from

$$((\Phi\Phi^*)^{-1/2}\Phi^*)^*(\Phi\Phi^*)^{-1/2}\Phi^* = I_n$$

and the definition of a Parseval frame. □

However, various properties of the frame such as the direction of the frame vectors is destroyed during the process. A much more careful procedure is to just scale each frame vector to generate a Parseval frame, which can also be regarded as preconditioning by a diagonal matrix. Characterizing conditions – also of geometric type – for a frame to be scalable in this sense are contained in [17].

If no frame is given, but only the dimension of the space n, and the number of frame vectors m, there are also explicit algorithms to construct a corresponding tight frame. A specific algorithm which accomplishes this goal which is of special relevance to this paper as the frames constructed are particularly sparse will be described in Section 3.1

3 Sparse Frames

In this section, we will study our guiding problem in the most general context: How sparse can a frame be when the embedding and ambient dimensions n and m are given? We will ask this question with regards to the following definition of a sparse frame:

Definition 3.1 Let $(e_j)_{j=1}^n$ be an orthonormal basis for $\mathbb{R}^n$. Then a frame $(\varphi_i)_{i=1}^m$ for $\mathbb{R}^n$ is called *k-sparse* with respect to $(e_j)_{j=1}^n$, if, for each $i \in \{1,\ldots,m\}$, there exists $J_i \subseteq \{1,\ldots,n\}$ such that

$$\varphi_i \in \operatorname{span}\{e_j : j \in J_i\}$$

and

$$\sum_{i=1}^n |J_i| = k.$$

Note that according to this definition, a frame being k-sparse is the same as the associated frame matrix having only k non-vanishing entries, hence this definition is in line with (1.3). The main goal of this section will be to find *optimally sparse* frames in the sense of the following definition.

Definition 3.2 Let $\mathscr{F}$ be the class of all frames for $\mathbb{R}^n$ consisting of m frame vectors, let $(\varphi_i)_{i=1}^m \in \mathscr{F}$, and let $(e_j)_{j=1}^n$ be an orthonormal basis for $\mathbb{R}^n$. Then $(\varphi_i)_{i=1}^m$ is called *optimally sparse in $\mathscr{F}$ with respect to $(e_j)_{j=1}^n$*, if $(\varphi_i)_{i=1}^m$ is k_1-sparse with respect to $(e_j)_{j=1}^n$ and there does not exist a frame $(\psi_i)_{i=1}^m \in \mathscr{F}$ which is k_2-sparse with respect to $(e_j)_{j=1}^n$ with $k_2 < k_1$.

Without any additional constraints, an optimally sparse frame will always be given by the canonical basis extended by zero vectors. However, since the first n frame coefficients of a vector will carry all the information, the frame representation cannot be called redundant. A way to circumvent this issue is to require the frame vectors to be unit norm, i.e. $\|\varphi_i\|_2 = 1$. This normalization ensures that the frame coefficients carry equal information. Even then an optimally sparse frame is easily found by extending the canonical basis by $m-n$ copies of the first (or any other) basis vector. The resulting frame, however, has another drawback: Its redundancy is distributed very unevenly. In the direction of the repeated frame vector it is indeed very redundant, whereas in all other directions the frame is not redundant at all. The spatial distribution of the redundancy is studied by Bodmann et al. in [2], and the authors show that this is reflected in the spectrum of the frame operator. Consequently, we will fix the spectrum of the frame operator as an additional constraint.

Arguably the most natural case is that the desired redundancy is evenly distributed, which corresponds to a spectrum consisting of a single point and hence a tight frame. Thus our main focus in this survey article will be on the class $\mathscr{F}(n,m)$ of unit norm, tight frames consisting of m vectors in n dimensions. Remark 3.8 will briefly discuss extensions to more general spectra.

3.1 Spectral Tetris – an Algorithm to Construct Sparse Tight Frames

Frames in $\mathscr{F}(n,m)$ are characterised by having a frame matrix with orthogonal rows with norm $\sqrt{m/n}$ and columns with norm one. Constructing a sparse unit norm, tight frame hence boils down to satisfy these constraints using as few entries of the matrix as possible and setting the rest to zero. The objective of the Spectral tetris algorithm, as introduced in [5] by Casazza et al., is to accomplish these goals in a greedy fashion, determining the frame matrix entries subsequently in a recursive way. Its only requirement is that $m/n \geq 2$; for such dimension pairings it constructs a unit-norm tight frame $(\varphi_i)_{i=1}^m$ for $\mathbb{R}^n$. The frame bound then automatically equals m/n. The detailed steps are provided in Algorithm 1. We remark that an extension to arbitrary spectra of the frame operator is described in [3].

By construction, the frames resulting from Algorithm 1 will always be rather sparse. Indeed, the two cases which the algorithm distinguishes in the *if-else* statement correspond to adding a column with two non-vanishing entries or to adding a

Algorithm 1 STTF: Spectral Tetris for Tight Frames

Input: (i) Dimension $n \in \mathbb{N}$, (ii) Number of frame elements $m \in \mathbb{N}$ with $m/n \geq 2$, (iii) An ONB $(e_j)_{j=1}^n$.
Output: Frame STTF(n,m).

Initialize: $i \leftarrow 1$, $\lambda_1 \leftarrow m/n$
for $j = 1 : n$ **do**
  $\lambda_{j+1} \leftarrow m/n$.
  while $\lambda_j \neq 0$ **do**
    if $\lambda_j < 1$ **then**
      $\varphi_i \leftarrow \sqrt{\frac{\lambda_j}{2}} \cdot e_j + \sqrt{1 - \frac{\lambda_j}{2}} \cdot e_{j+1}$.
      $\varphi_{i+1} \leftarrow \sqrt{\frac{\lambda_j}{2}} \cdot e_j - \sqrt{1 - \frac{\lambda_j}{2}} \cdot e_{j+1}$.
      $i \leftarrow i+2$.
      $\lambda_{j+1} \leftarrow \lambda_{j+1} - (2 - \lambda_j)$.
      $\lambda_j \leftarrow 0$.
    else
      $\varphi_i \leftarrow e_j$.
      $i \leftarrow i+1$.
      $\lambda_j \leftarrow \lambda_j - 1$.
    end if
  end while
end for
STTF$(n,m) \leftarrow \{\varphi_i\}_{i=1}^m$.

column with just a single non-vanishing entry. Thus in any case, the frame has sparsity at most $2m$, which should be considered small compared to the total number of mn matrix entries. The intrinsic sparsity of the algorithmic procedure is best visualized by the example Spectral Tetris frame given in (3.1) below.

Our goal for the remainder of this section is first to find a lower bound for the optimal sparsity within the class $\mathscr{F}(n,m)$ and then to show that the Spectral Tetris frames, when they exist, in fact obtain this lower bound. As we will see in the following subsection, a factor that determines the optimal sparsity and that is consequently a key proof ingredient is whether the dimension pairing allows for block decompositions.

3.2 Block Decompositions of Frame Matrices

Recall that Algorithm 1 returns a frame consisting only of 1-sparse and 2-sparse vectors. Hence the only way one can expect to find an even sparser frame would be to have fewer 2-sparse vectors than the output of Algorithm 1. So it is intuitively relevant to ask for the minimal number of 2-sparse vectors that a frame of given embedding and ambient dimensions must have. Spectral tetris frames contain at most $2n-2$ frame vectors which are 2-sparse, but can contain less. If this is the case, the frame matrix has a block decomposition of order at least 2 in the sense of the following definition.

Definition 3.3 Let $n,m > 0$, and let $(\varphi_i)_{i=1}^m$ be a frame for $\mathbb{R}^n$. Then we say that the frame matrix $\Phi = (\varphi_i)_{i=1}^m$ has *block decomposition of order* ν, if there exists a partition $\{1,\ldots,m\} = I_1 \cup \ldots \cup I_\nu$ such that, for any $i_1 \in I_{\ell_1}$ and $i_2 \in I_{\ell_2}$ with $\ell_1 \neq \ell_2$, we have $\operatorname{supp}\varphi_{i_1} \cap \operatorname{supp}\varphi_{i_2} = \emptyset$ and ν is maximal.

The following result now connects the block decomposition order of a frame matrix with the greatest common divisor of the dimension and the number of frame vectors.

Proposition 3.4 ([7]) *Let $m > n > 0$ and $\Phi = (\varphi_i)_{i=1}^m \in \mathscr{F}(n,m)$. Then the frame matrix Φ has block decomposition of order at most* $\gcd(n,m)$.

Proof Assume $(\varphi_i)_{i=1}^m \in \mathscr{F}(n,m)$ has block decomposition of order ν, and let a corresponding partition be given by $\{1,\dots,m\} = I_1 \cup \ldots \cup I_\nu$. For each $\ell = 1,\dots,\nu$, let S_ℓ be the common support set of the vectors $(\varphi_i)_{i\in I_\ell}$. In other words, we have $k \in S_\ell$ if and only if $k \in \operatorname{supp}\varphi_i$ for some $i \in I_\ell$. Further, let r_k denote the kth row of the frame matrix of $(\varphi_i)_{i=1}^m$; note that $\|r_k\|^2 = \frac{m}{n}$ as the frame is tight. Then $S_1,\dots,S_\nu$ is a partition of $\{1,\dots,n\}$ and, for every $\ell = 1,\dots,\nu$, we have by the fact that $(\varphi_i)_{i=1}^m$ consists of unit norm vectors and by our choice of I_ℓ and S_ℓ that

$$\#I_\ell = \sum_{k\in I_\ell} \|\varphi_k\|^2 = \sum_{k\in S_\ell} \|r_k\|^2 = \sum_{k\in S_\ell} \frac{m}{n} = \#S_\ell \frac{m}{n}.$$

The second equality holds since we, after permutation of the columns, can write the frame matrix of $(\varphi_i)_{i=1}^m$ as $\Phi = [\Phi_1,\dots,\Phi_\nu]$, where Φ_ℓ has zero entries except on the rows indexed by I_ℓ and the columns indexed by S_ℓ, for $\ell = 1,\dots,\nu$. As each $\#I_\ell$ is an integer, one must have $\#S_\ell \geq \frac{n}{\gcd(n,m)}$ for all ℓ, and hence we obtain $\nu \leq \frac{n}{\min_\ell \#S_\ell} \leq \gcd(n,m)$. □

3.3 A Lower Bound for Sparsity within $\mathscr{F}(n,m)$

The main result of the section, Theorem 3.5, is a lower bound on the achievable sparsity of frames in the class $\mathscr{F}(n,m)$. With Theorem 3.6 in the next section, we will see that this lower bound is actually realizable for any $n,m \in \mathbb{N}$ in the range $m \geq 2n$.

Theorem 3.5 ([7]) *Let $m > n > 0$. Suppose that $(\varphi_i)_{i=1}^m \in \mathscr{F}(n,m)$. Then*

$$\|\Phi\|_0 \geq m + 2(n - \gcd(n,m)),$$

where Φ is the frame matrix of $(\varphi_i)_{i=1}^m$ with respect to any orthonormal basis.

Proof Let Φ denote the frame matrix of a frame in $\mathscr{F}(n,m)$ with respect to a fixed orthonormal basis. For the sake of brevity, in the sequel we will use the phrase that two rows of Φ *have overlap of size k*, if the intersection of their supports is a set of size k. Note that, since the rows of Φ are orthogonal, it is not possible that two rows of Φ have overlap of size one.

We first consider the case where $\gcd(n,m) = 1$. Pick an arbitrary row r_1 of Φ. Since, by Proposition 3.4, Φ has block decomposition of order one, there exists a row r_2 whose overlap with r_1 is of size two or greater. Similarly, there has to exist a row different from r_1 and r_2 which has overlap of size two or greater with either r_1 or r_2. Iterating this procedure will provide an order $r_1, r_2, \ldots r_n$ such that, for each row r_j, there exists some $k < j$ such that r_j has overlap of size two or greater with r_k. Since all columns in Φ are of norm one, for each column v, there exists a minimal j for which the r_jth entry of the vector v is non-zero. This yields m non-zero entries in Φ. In addition, each row r_2 through r_n has at least two non-zero entries coming from the overlap, which are different from the just accounted for m entries, since these entries cannot be the non-zero entries of minimal index of a column due to the overlap with a previous row. This sums up to a total of at least $2(n-1)$ non-zero coefficients. Consequently, the frame matrix has at least $m+2(n-1)$ non-zero entries.

We now consider the case where $\gcd(n,m) = \nu > 1$. By Proposition 3.4, Φ has block decomposition of order at most ν. Performing the same construction as above, we see that there exist at most ν rows r_j (including the first one) which do not have overlap with a row r_k for $k < j$. Thus the frame matrix Φ must at least contain $m + 2(n-\nu)$ non-zero entries. □

3.4 Optimally Sparse Unit Norm Tight Frames

Having set the benchmark, we now prove that frames constructed by Spectral Tetris in fact meet the optimal sparsity rate. For this, we would like to remind the reader that the tight Spectral Tetris frame as constructed by Algorithm 1 is denoted by STTF(n,m).

Theorem 3.6 ([7]) *Let $m \geq 2n > 0$. Let $(e_j)_{j=1}^n$ be an orthonormal basis for $\mathbb{R}^n$. Then the frame STTF(n,m) constructed using $(e_j)_{j=1}^n$ is optimally sparse in $\mathscr{F}(n,m)$ with respect to $(e_j)_{j=1}^n$. That is, if Φ is the frame matrix of STTF(n,m) with respect to $(e_j)_{j=1}^n$, then*

$$\|\Phi\|_0 = m + 2(n - \gcd(n,m)).$$

Proof Let $(\varphi_i)_{i=1}^m$ be the frame STTF(n,m), and let Φ be the frame matrix of $(\varphi_i)_{i=1}^m$ with respect to $(e_j)_{j=1}^n$. We start by showing that Φ has block decomposition of order $\nu = \gcd(n,m)$. For this, we set $n_0 = m_0 = 0$ and $n_j = j\frac{n}{\nu}$, $m_j = j\frac{m}{\nu}$, $1 \leq j \leq \nu$; in particular this entails $k_\nu = n$ and $m_\nu = m$.

As $\frac{m}{n} = \frac{m_1}{n_1}$, the first n_1 entries of the first m_1 coefficient vectors resulting from STTF(n_1,m_1) and STTF(n,m) coincide; indeed, the corresponding steps are identical. Continuing the computation of STTF(n,m) will set the remaining entries of the first m_1 vectors and also the first n_1 entries of the remaining vectors to zero. Thus, any of the first n_1 vectors has disjoint support from any of the vectors constructed later on. Repeating this argument for n_2 until n_ν, we obtain that Φ has a block decomposition of order ν; the corresponding partition of the frame vectors being

$$\bigcup_{i=1}^{\nu} \{\varphi_{m_{i-1}+1}, \ldots, \varphi_{m_i}\}.$$

To compute the number of non-zero entries in Φ, we let $i \in \{1,\ldots,\nu\}$ be arbitrarily fixed and compute the number of non-zero entries of the vectors $\varphi_{m_{i-1}+1},\ldots,\varphi_{m_i}$. Spectral Tetris ensures that the support of each of the rows $n_{i-1}+1$ up to $n_i - 1$

intersects the support of the subsequent row on a set of size two, as otherwise an additional block would be created. Thus, there exist $2(n_i - n_{i-1} - 1)$ frame vectors with two non-zero entries. The remaining $(m_i - m_{i-1}) - 2(n_i - n_{i-1} - 1)$ frame vectors will have only one entry, yielding a total number of $(m_i - m_{i-1}) + 2(n_i - n_{i-1} - 1)$ non-zero entries in the vectors $\varphi_{m_{i-1}+1}, \ldots, \varphi_{m_i}$.

Summarizing, the total number of non-zero entries in the frame vectors of $(\varphi_i)_{i=1}^m$ is

$$\begin{aligned}\sum_{i=1}^{\nu}(m_i - m_{i-1}) + 2(n_i - n_{i-1} - 1) &= \left(\sum_{i=1}^{\nu}(m_i - m_{i-1})\right) + 2\left(n_\nu - \left(\sum_{i=1}^{\nu} 1\right)\right) \\ &= m + 2(n - \nu),\end{aligned}$$

which, by Theorem 3.5, is the optimal sparsity. □

The following example shows that an optimally sparse frame from $\mathscr{F}(n,m)$ is, in general, not unique.

Example 3.7 Let $n = 4$ and $m = 9$. Then, by Theorem 3.5, the optimal sparsity is $9 + 2(4-1) = 15$. The following matrices are frame matrices with respect to a given orthonormal basis of two different unit-norm tight frames in $\mathbb{R}^4$:

$$\Phi_1 = \begin{bmatrix} 1 & 1 & \sqrt{\frac{1}{8}} & \sqrt{\frac{1}{8}} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{\frac{7}{8}} & -\sqrt{\frac{7}{8}} & \sqrt{\frac{1}{4}} & \sqrt{\frac{1}{4}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{\frac{3}{4}} & -\sqrt{\frac{3}{4}} & \sqrt{\frac{3}{8}} & \sqrt{\frac{3}{8}} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{\frac{5}{8}} & -\sqrt{\frac{5}{8}} & 1 \end{bmatrix} \tag{3.1}$$

and

$$\Phi_2 = \begin{bmatrix} 1 & \sqrt{\frac{5}{8}} & \sqrt{\frac{5}{8}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sqrt{\frac{3}{8}} & -\sqrt{\frac{3}{8}} & \sqrt{\frac{3}{8}} & \sqrt{\frac{3}{8}} & \sqrt{\frac{3}{8}} & \sqrt{\frac{3}{8}} & 0 & 0 \\ 0 & 0 & 0 & \sqrt{\frac{5}{8}} & -\sqrt{\frac{5}{8}} & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sqrt{\frac{5}{8}} & -\sqrt{\frac{5}{8}} & 0 & 1 \end{bmatrix}.$$

Clearly, both frame matrices satisfy $\|\Phi_1\|_0 = \|\Phi_2\|_0 = 15$, hence the corresponding frames are optimally sparse in $\mathscr{F}(4,9)$. We remark that Φ_1 is the frame matrix of STTF$(4,9)$.

Remark 3.8 Tight frames have the special property that their spectrum uniquely defines the frame operator. While it is shown in [7] that given a diagonal frame operator, Spectral Tetris for Frames, a variant of Algorithm 1, always constructs the optimally sparse frame associated to that operator, there may exist a frame with a non-diagonal frame operator of the same spectrum which is sparser. Hence in such cases, Spectral Tetris does not always find the sparsest frame of a given spectrum.

Remark 3.9 For redundancy less than 2, the picture is less clear. It has been shown [6, 18] that the Spectral Tetris algorithm works even in certain cases where $m < n$. However, to our knowledge, a systematic analysis of optimal sparsity for low redundancy has not been performed.

4 Sparse Dual Frames

In this section we will take a different viewpoint regarding sparsity, and we now ask how sparse a dual frame of a given, fixed frame can be. This viewpoint is motivated by the fact that in many cases the decomposition frame $(\varphi_i)_{i=1}^m$ is given by the application at hand, e.g., by the way of measuring the data. In these situations, one is then interested in choosing a good dual frame of $(\varphi_i)_{i=1}^m$ for the reconstruction process. Notice again how the redundancy of frames plays a key role here since in the non-redundant case we would only have one option for exact reconstruction.

Again, the goal will be to achieve sparsity in the sense of Definitions 3.1, with the goal of allowing for efficient matrix-vector multiplication, just this time for the dual frame (Note that in contrast to the previous section, the basis used for representing the dual frame is the same as for the frame at hand and hence fixed). That is, we are interested in the following minimization problem:

$$\min \|\Psi\|_0 \quad \text{s.t.} \quad \Phi\Psi^* = I_n, \tag{4.1}$$

where the frame Φ is given. In line with Definition 3.2, we name the solutions of (4.1) *optimally sparse duals*; note that as for optimally sparse frames, these solutions are, in general, not unique. In the following two subsections we will focus on analyzing the possible values of the objective function $\|\Psi\|_0$ in (4.1), and in the final subsection we will comment on the problem of finding the minimizers.

4.1 Optimal Sparsity of Dual Frames

In this subsection we investigate the possible sparsity levels in the set of all dual frames. We start with a trivial upper bound.

Lemma 4.1 *Suppose Φ is a frame for $\mathbb{R}^n$. Then there exists a dual frame Ψ of Φ with*

$$\|\Psi\|_0 \le n^2.$$

Proof Let $J \subset \{1,2,\ldots,m\}$ be such that $\#J = n$ and the corresponding frame vectors $(\varphi_i)_{i\in J}$ are linearly independent. Such a set always exists, as in order to form a frame, the columns of Φ must span $\mathbb{R}^n$. Let $(\psi_i)_{i\in J}$ be the unique (bi-orthogonal) dual of $(\varphi_i)_{i\in J}$, and set $\psi_i = 0$ for $i \notin J$. Then we obviously have $\|\Psi\|_0 \le n^2$. □

Without additional assumptions on the frame, only the trivial lower bound $\|\Psi\|_0 \ge n$ on the dual frame sparsity can be established; this bound is achieved, e.g., for frames containing the canonical basis vectors. To precisely determine the optimal sparsity in the set of all duals of a given frame, we thus need to take into account further properties of the frame. In particular, the spark of the frame matrix plays an important role. Recall that the spark of a matrix $\Phi \in \mathbb{R}^{n\times m}$ is defined as the smallest number of linear dependent columns of Φ and denoted by spark (Φ). For an invertible $n\times n$ matrix Φ, one sets spark $(\Phi) = n+1$. In fact, we will need the following refined version of the spark.

Definition 4.2 Let $\Phi \in \mathbb{R}^{n\times m}$. Then $\operatorname{spark}_j(\Phi)$ denotes the smallest number of linearly dependent columns in the $(n-1)\times m$ submatrix $\Phi^{(j)}$ of Φ with the jth row deleted such that the corresponding columns in Φ are linear independent.

We can now state the main result of this subsection.

Theorem 4.3 ([14]) *Suppose Φ is a frame for $\mathbb{R}^n$. Then the optimally sparse dual frame Ψ of Φ satisfies*

$$\|\Psi\|_0 = \sum_{j=1}^{n} \operatorname{spark}_j(\Phi).$$

Proof Let Ψ be an optimally sparse dual of Φ. Fix $j \in \{1,\ldots,n\}$, and let $\varphi_k^{(j)}$ denote the kth column of $\Phi^{(j)}$. Since Φ and Ψ are assumed to be dual frames, we have that $\Phi^{(j)}\left(\psi^j\right)^* = 0_{n-1}$ to hold, where ψ^j denotes the jth row of Ψ. This shows that $\left(\varphi_k^{(j)}\right)_{k\in \operatorname{supp}\psi^j}$ must be linearly dependent. On the other hand, the frame vectors $(\varphi_k)_{k\in\operatorname{supp}\psi^j}$ must be linearly independent, as otherwise one of these columns would be a linear combination of the others, which would allow for the construction of $\tilde{\psi}^j$ with $\operatorname{supp}\tilde{\psi}^j \subsetneq \operatorname{supp}\psi^j$ such that $\Phi\left(\tilde{\psi}^j\right)^* = e_j$. This in turn would imply that the frame, whose frame matrix is obtained from Ψ by replacing the row ψ^j by $\tilde{\psi}^j$ is also a dual frame of Φ, so Ψ is not the optimally sparse dual, contradicting our assumption. Therefore we obtain $|\operatorname{supp}\psi^j| \geq \operatorname{spark}_j(\Phi)$ which, in turn, implies that $\|\Psi\|_0 \geq \sum_{j=1}^n \operatorname{spark}_j(\Phi)$.

To complete the proof we need to show existence of a dual frame that meets this lower bound. To that intent, fix $j \in \{1,\ldots,n\}$, and let S be a set of size $\operatorname{spark}_j(\Phi)$ such that $(\varphi_k)_{k\in S}$ is a set of linearly independent columns of Φ such that the corresponding columns of $\Phi^{(j)}$ are linearly dependent. That is, there exist $(\lambda_k)_{k\in S}$ such that $\sum_{k\in S}\lambda_k(\varphi^{(j)})_k = 0$, but $(\sum_{k\in S}\lambda_k\varphi_k)_j = a \neq 0$. This motivates to define ψ^j via

$$\psi_k^j = \begin{cases} \frac{\lambda_k}{a} & \text{if } k \in S,\\ 0 & \text{if } k \in \{1,\ldots,m\}\setminus S,\end{cases}$$

as this definition yields $\Phi(\psi^j)^* = e_j$. Therefore, the matrix Ψ with rows ψ^j, $j \in \{1,\ldots,n\}$, corresponds to a dual frame which is $\sum_{j=1}^n \operatorname{spark}_j(\Phi)$-sparse. □

By definition we have $\operatorname{spark}_j(\Phi) \geq \operatorname{spark}(\Phi^{(j)})$ for every $j = 1,\ldots,n$. Hence, we immediately have the following useful corollary of Theorem 4.3.

Corollary 4.4 *Suppose Φ is a frame for $\mathbb{R}^n$. Then any dual frame Ψ of Φ satisfies*

$$\|\Psi\|_0 \geq \sum_{j=1}^{n} \operatorname{spark}(\Phi^{(j)}).$$

4.2 n^2-Sparse Duals

By Lemma 4.1 we know that it is always possible to find a dual frame with sparsity level n^2. The result below states that for a large class of frames this is actually the best, or rather the sparsest, one can achieve. Recall that an $n \times m$ matrix is said to be in *general position* if any sub-collection of n (column) vectors is linearly independent, that is, if any $n \times n$ submatrix is invertible. Such matrices are sometimes called *full spark frames*, such as in [1], since their spark is maximal, i.e., $n+1$.

Theorem 4.5 ([14]) *Suppose* Φ *is a frame for* $\mathbb{R}^n$ *such that the submatrix* $\Phi^{(j)}$ *is in general position for every* $j = 1, \ldots, n$*. Then any dual frame* Ψ *of* Φ *satisfies*

$$\|\Psi\|_0 \geq n^2. \tag{4.2}$$

In particular, the optimally sparse dual satisfies $\|\Psi\|_0 = n^2$

Proof Since $\Phi^{(j)}$ is in general position, we have $\operatorname{spark}(\Phi^{(j)}) = n$ for each $j = 1, \ldots, n$. By Corollary 4.4 we immediately have (4.2), and then, by Lemma 4.1, that the optimally sparse satisfies $\|\Psi\|_0 = n^2$. □

We illustrate this result with a number of examples of frames which are well-known to be in general position, and which thus do not allow for dual frames with less than n^2 non-vanishing entries.

Example 4.6 For any $n, m \in \mathbb{N}$ with $m \geq n$, let $a_i > 0$, $i = 1, \ldots, m$ with $a_i \neq a_j$ for all $i \neq j$, and let $b_j \in \mathbb{R}$, $j = 1, \ldots, n$ with $b_j \neq b_i$, $j \neq i$. Generalized Vandermonde frames are defined by:

$$\Phi = \begin{bmatrix} a_1^{b_1} & a_1^{b_2} & \cdots & a_1^{b_m} \\ \vdots & \vdots & & \vdots \\ a_n^{b_1} & a_n^{b_2} & \cdots & a_n^{b_m} \end{bmatrix}.$$

It is well-known, see e.g., [13, §8.1], that the submatrix $\Phi^{(j)}$ is in general position for every $j = 1, \ldots, m$, which, by Theorem 4.5, implies that optimally sparse duals of Φ are n^2-sparse.

As previously mentioned the results in this paper also hold for complex frames. The next two examples are frames for $\mathbb{C}^n$ whose sparsest dual has sparsity n^2. Notice that we here deviate from our policy to restrict to real frames.

Example 4.7 Let $n \in \mathbb{N}$, and let m be prime. Let Φ be a partial FFT matrix of size $n \times m$, that is, Φ is constructed by picking n rows of an $m \times m$ DFT matrix at random. We remark that Φ is a tight frame. Moreover, any $\Phi^{(j)}$ is in general position, as the determinant of any $(n-1) \times (n-1)$ submatrix of Φ is non-zero which is a consequence of Chebotarev theorem about roots of unity stating that any minor of an $m \times m$ DFT matrix is non-zero whenever m is prime [19, 20]. Our conclusion is again that the sparsest dual frame Ψ of Φ satisfies $\|\Psi\|_0 = n^2$.

Example 4.8 For $n \in \mathbb{N}$ prime and $m = n^2$, Krahmer et al. showed in [15] that for almost every $v \in \mathbb{C}^n$, the Gabor frame generated by v has a frame matrix with no zero minors. For this reason, the optimally sparse duals of such Gabor frames have sparsity n^2.

In fact, the property of having a sparsest dual with sparsity n^2 is a *very generic* property in the sense that the set $\mathcal{N}(n,m)$ of all frames of m vectors in $\mathbb{R}^n$ whose sparsest dual has n^2 non-zero entries contains an open, dense subset in $\mathbb{R}^{n\times m}$ and has a compliment of measure zero. Let $\mathscr{P}(n,m)$ be the the set of all frames Φ which satisfy spark $(\Phi^{(j)}) = n$ for all $j = 1,\ldots,n$. By Corollary 4.4 we have that $\mathscr{P}(n,m) \subset \mathcal{N}(n,m)$. We are now ready to state the result saying that "most" frame has a sparsest dual with sparsity level n^2.

Lemma 4.9 *Suppose $m \geq n$. Then the set $\mathscr{P}(n,m)$ is open and dense in $\mathbb{R}^{n\times m}$, and $\mathscr{P}(n,m)^c$ is of measure zero.*

Proof Note that $\Phi = [x_{k,\ell}]_{k\in\{1,\ldots,n\},\ell\in\{1,\ldots,m\}} \in \mathscr{P}(n,m)$ if and only if Φ has full-rank and the polynomials in $x_{k,\ell}$ given by $\det\left([\Phi^{(j)}]_I\right)$ are non-zero for each $j \in \{1,\ldots,m\}$ and $I \in S$, where S is the collection of all subsets of $\{1,\ldots,m\}$ of size $n-1$. Here $[\Phi^{(j)}]_I \in \mathbb{R}^{(n-1)\times(n-1)}$ denotes the matrix $\Phi^{(j)}$ restricted to the columns in the index set I. This shows that $\mathscr{P}(n,m)$ is open in the Zariski topology. Since the set $\mathscr{P}(n,m)$ is non-empty by Example 4.6, it is thereby open and dense in the standard topology, see [1]. Finally, since $\mathscr{P}(n,m)^c$ is a proper subset and closed in the Zariski topology, it is of measure zero. □

By Lemma 4.9 we see that any frame of m vectors in $\mathbb{R}^n$ is arbitrarily close to a frame in $\mathcal{N}(n,m)$.

Theorem 4.10 ([14]) *Every frame is arbitrarily close to a frame whose sparsest dual Ψ satisfies $\|\Psi\|_0 = n^2$.*

Another consequence of Lemma 4.9 is that for many randomly generated frames, the sparsest dual has sparsity level $\|\Psi\|_0 = n^2$. As an example, this holds when the entries of Φ are drawn independently at random from a standard normal distribution or when the frame obtained by a small Gaussian random perturbation of a given $n \times m$ matrix.

5 Discussion and Future Directions

Comparing the problems discussed in Sections 3 and 4, one observes that the corresponding optimally sparse solutions behave very differently. For optimally sparse frames, a large fraction of the entries vanishes (only between n and $2n$ entries will not vanish), hence there is a limited number of degrees of freedom and the resulting frames are structurally very similar. For optimally sparse duals, in contrast, the sparsity can vastly vary. For frames in general position, the minimal number of non-vanishing entries is n^2 and hence rather large. As a consequence, there is a larger number of degrees of freedom that allows for structurally very different solutions.

For example, the n^2 non-vanishing entries could be all in the same column or evenly distributed over all columns.

Consequently, potential future research directions relating to the two problems are rather different. For sparse frames, we understand the cases presented above where $m \geq n$ quite well. But as mentioned in Remark 3.9, the low redundancy cases are open. Regarding sparse duals, we can completely characterize the achievable sparsity levels, but the question remains which of the many very different optimally sparse duals are most desired from the computational viewpoint. As mentioned above, a dual frame with all n^2 non-vanishing entries in the same columns discards all of the redundant information, so intuitively it is less desirable than distributing the entries evenly over the columns. Hence, it is important for a complete understanding to quantify this preference for one optimally sparse dual frame over the other in order to decide between two duals in case such an intuitive judgment is not possible. As the number of possible optimally sparse duals is very large, a satisfactory answer to this question must also involve algorithms to find duals which are at least close to the desired optimum. Furthermore, it would be interesting to find concrete examples of frame classes that allow for duals which are considerably sparser than in the generic case.

Acknowledgements The second author acknowledges support by the Einstein Foundation Berlin, by Deutsche Forschungsgemeinschaft (DFG) Grants SPP-1324 KU 1446/13 and KU 1446/14, and by the DFG Research Center MATHEON "Mathematics for key technologies" in Berlin.

References

1. B. Alexeev, J. Cahill, D. G. Mixon, Full spark frames, preprint
2. B. G. Bodmann, P. Casazza, and G. Kutyniok, A quantitative notion of redundancy for finite frames, Appl. Comput. Harmon. Anal., 30, 348–362 (2011)
3. R. Calderbank, P. Casazza, A. Heinecke, G. Kutyniok, and A. Pezeshki, Sparse fusion frames: existence and construction, Adv. Comput. Math., 35, 1–31 (2011)
4. E. J. Candès, J. Romberg, and T. Tao, Stable signal recovery from incomplete and inaccurate measurements, Comm. Pure Appl. Math., 59, 1207–1223 (2006)
5. P. G. Casazza, M. Fickus, D. G. Mixon, Y. Wang, and Z. Zhou, Constructing tight fusion frames, Appl. Comput. Harmon. Anal., 30, 175–187 (2011)
6. P. Casazza, A. Heinecke, K. Kornelson, Y. Wang, and Z. Zhou, Necessary and sufficient conditions to perform Spectral Tetris, preprint
7. P. G. Casazza, A. Heinecke, F. Krahmer, and G. Kutyniok, Optimally Sparse Frames, IEEE Trans. Inform. Theory, 57, 7279–7287 (2011)
8. P. G. Casazza and G. Kutyniok, eds., Finite Frames: Theory and Applications. Birkhäuser, Boston (2012)
9. I. Daubechies, A. Grossman, and Y. Meyer, Painless nonorthogonal expansions, J. Math. Phys., 27, 1271–1283 (1985)
10. D.L. Donoho, Compressed sensing, IEEE Trans. Inform. Theory, 52, 1289–1306 (2006)
11. R. Duffin and A. Schaeffer, A class of nonharmonic Fourier series, Trans. Amer. Math. Soc., 72, 341–366 (1952)
12. Y. C. Eldar and G. Kutyniok, eds., Compressed Sensing: Theory and Applications. Cambridge University Press, Cambridge (2012)
13. F. R. Gantmacher, Applications of the theory of matrices, Translated by J. L. Brenner, with the assistance of D. W. Bushaw and S. Evanusa. Interscience Publishers, Inc., New York (1959)
14. F. Krahmer, G. Kutyniok, and J. Lemvig, Sparsity and spectral properties of dual frames, preprint
15. F. Krahmer, G. Pfander, and P. Rashkov, Uncertainty in time-frequency representations on finite abelian groups and applications, Appl. Comput. Harmon. Anal., 25, 209–225 (2008)

16. G. Kutyniok, Theory and Applications of Compressed Sensing, preprint
17. G. Kutyniok, K. A. Okoudjou, F. Philipp, and E. K. Tuley, Scalable Frames, preprint
18. J. Lemvig, C. Miller, and K. A. Okoudjou, Prime tight frames, preprint
19. F. Pakovich, A remark on the Chebotarev theorem about roots of unity, Integers, 7, A18 (2007)
20. P. Stevenhagen and H. W. Lenstra, Chebotarëv and his density theorem, Math. Intelligencer, 18, 26–37 (1996)