SCALABLE PROBABILISTIC FRAMES

CLARE WICKMAN LAU AND KASSO A. OKOUDJOU

Abstract. We consider the problem of rescaling the lengths of a finite frame thereby transforming it into a tight one. Such frames are called scalable and have received a lot of attention in recent years. In this note we investigate the question in terms of probabilistic frames and give conditions under which a (discrete) probabilistic frame is scalable.

1. Introduction

Finite frame theory is by now a very-well established research area owing in part to the fundamental role frames’s inherent redundancy plays in many applications [1]. For example, frames provide an intuitive framework for describing and solving problems in coding theory, analog-to-digital quantization theory, sparse representation, and compressive sensing, and more generally they have proven useful in work on signal processing for telecommunications. We recall that \( \Phi = \{ \varphi_i \}_{i=1}^N \subset \mathbb{R}^d \) is a frame if there exist frame bounds \( 0 < A \leq B < \infty \) such that

\[
\forall x \in \mathbb{R}, \quad A \| x \|^2 \leq \sum_{i=1}^N \langle x, \varphi_i \rangle^2 \leq B \| x \|^2.
\]

Developed in a series of papers ([2, 3, 4]), probabilistic frames are one way to generalize finite frames. In the simplest example, each finite frame can be used to build a probabilistic frame. Taking \( \Phi \) above, let \( \{ \alpha_i \}_{i=1}^N \) be a set of positive real numbers satisfying \( \sum_{i=1}^N \alpha_i = 1 \). Then the canonical \( \alpha \)-weighted probabilistic frame for \( \Phi \) is \( \mu_{\Phi, \alpha} \) given by

\[
d\mu_{\Phi, \alpha}(x) = \sum_{i=1}^N \alpha_i \delta_{\varphi_i}(x).
\]

More generally, a probabilistic frame \( \mu \) for \( \mathbb{R}^d \) is a probability measure on \( \mathbb{R}^d \) for which there exist constants \( 0 < A \leq B < \infty \) such that for all \( x \in \mathbb{R}^d \),

\[
A \| x \|^2 \leq \int_{\mathbb{R}^d} \langle x, y \rangle^2 d\mu(y) \leq B \| x \|^2.
\]
This amounts to a restriction on the covariance of the probability measure. Probabilistic frames are related to statistical shape analysis, as detailed in [3], and they are linked to the classical problem of estimating the population covariance from a sample [4, 5]. In the probabilistic setting, other questions from frame theory can be explored, such as the scaling problem we consider in this paper.

The most natural space to explore probabilistic frames is the Wasserstein space of probability measures with finite second moment, a metric space with distance defined by the concept of optimal transport. A probability measure \( \mu \) on \( \mathbb{R}^d \) is an element of \( \mathcal{P}_2(\mathbb{R}^d) \), the space of probability measures with finite second moment, if it satisfies:

\[
M_2^2(\mu) := \int_{\mathbb{R}^d} \|x\|^2 d\mu(x) < \infty
\]

The support of a probability measure \( \mu \) on \( \mathbb{R}^d \) is the set \( \text{supp}(\mu) \) given by:

\[
\{ x \in \mathbb{R}^d \text{ s.t. for all open sets } U_x \text{ containing } x, \mu(U_x) > 0 \}.
\]

By [2, Theorem 5], a probability measure \( \mu \) on \( \mathbb{R}^d \) is a probabilistic frame if and only if it has finite second moment, and the linear span of its support is \( \mathbb{R}^d \). One can restate this in terms of the probabilistic frame operator, which, given a measure \( \mu \in \mathcal{P}_2(\mathbb{R}^d) \), is the operator \( S_\mu \) which for all \( y \in \mathbb{R}^d \) satisfies:

\[
S_\mu y = \int_{\mathbb{R}^d} \langle x , y \rangle x d\mu(x).
\]

Clearly, \( S_\mu \) may be equated with its matrix representation \( \int_{\mathbb{R}^d} xx^\top d\mu(x) \), and then the requirement that the support of \( \mu \) span \( \mathbb{R}^d \) is the same as requiring that this matrix be positive definite.

One of the most useful metrics on \( \mathcal{P}_2(\mathbb{R}^d) \) is the Wasserstein distance; it metrizes the weak convergence on the space. The Wasserstein distance between two probability measures \( \mu \) and \( \nu \) on \( \mathbb{R}^d \) is:

\[
W_2^2(\mu, \nu) := \inf _{\gamma} \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} \|x - y\|^2 d\gamma(x, y) : \gamma \in \Gamma(\mu, \nu) \right\},
\]

where \( \Gamma(\mu, \nu) \) is the set of all joint probability measures \( \gamma \) on \( \mathbb{R}^d \times \mathbb{R}^d \) such that for all \( A, B \subset \mathcal{B}(\mathbb{R}^d) \), \( \gamma(A \times \mathbb{R}^d) = \mu(A) \) and \( \gamma(\mathbb{R}^d \times B) = \nu(B) \). The search for the set of joint measures which induce the infimum is a variant of the Monge-Kantorovich problem. A joint distribution \( \gamma_0 \) which induces this infimum is called an optimal transport plan. In
the quadratic case, when $\mu$ and $\nu$ do not assign positive measure to isolated points, then

$$W_2^2(\mu, \nu) := \inf_T \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} \| x - T(x) \|^2 d\mu(x) : T_*\mu = \nu \right\},$$

where $T$ is a deterministic transport map (or deterministic coupling): i.e., for all $\nu$-integrable functions $\phi$,

$$\int_{\mathbb{R}^d} \phi(y) d\nu(y) = \int_{\mathbb{R}^d} \phi(T(x)) d\mu(x).$$

Equipped with the 2-Wasserstein distance, $P_2(\mathbb{R}^d)$ is a complete, separable metric space. Convergence in $P_2(\mathbb{R}^d)$ is the usual weak convergence of probability measures, combined with convergence of the second moments. Alternatively, one can view this convergence as simply enlarging the set of allowable test functions for convergence and write that a sequence of measures $\mu_n \in P_\rho(\mathbb{R}^d)$ is said to converge weakly to $\mu \in P_\rho(\mathbb{R}^d)$ if for all continuous functions $\phi$ with

$$|\phi(x)| \leq C(1 + \|x - x_0\|^p),$$

for some $C > 0$ and some $x_0 \in \mathbb{R}^d$,

$$\int_{\mathbb{R}^d} \phi(x) d\mu_n(x) \rightarrow \int_{\mathbb{R}^d} \phi(x) d\mu(x).$$

Several structural statements can be made about probabilistic frames as a subset of $P_2(\mathbb{R}^d)$. For brevity, let us denote the probabilistic frames for $\mathbb{R}^d$ by $PF(\mathbb{R}^d)$, and let $PF(A, B, \mathbb{R}^d)$ denote the set of probabilistic frames in $PF(\mathbb{R}^d)$ with upper frame bound less than or equal to $B$ and lower frame bound greater than or equal to $A$.

**Proposition 1.** Given finite $A, B > 0$, $PF(A, B, \mathbb{R}^d)$ is a nonempty, convex, closed subset of $P_2(\mathbb{R}^d)$.

**Proof.** The nonemptiness is clear: consider the space of nondegenerate, zero-mean Gaussian measures on $\mathbb{R}^d$ whose covariance matrices have maximum eigenvalue $B$ and minimum eigenvalue $A$. For the convexity: consider $\mu, \nu \in PF(A, B, \mathbb{R}^d)$, $\lambda \in [0, 1]$. Define $\mu_\lambda = (1 - \lambda)\mu + \lambda\nu$. Given $y \in \mathbb{R}^d$,

$$\int_{\mathbb{R}^d} \langle x, y \rangle^2 d\mu_\lambda(x) = (1 - \lambda) \int_{\mathbb{R}^d} \langle x, y \rangle^2 d\mu(x)$$

$$+ \lambda \int_{\mathbb{R}^d} \langle x, y \rangle^2 d\nu(x)$$

$$\geq A \|y\|^2$$
The upper bound is trivial.

Finally, for the closedness, let \( \{ \mu_n \} \) be a sequence in \( \text{PF}(A, B, \mathbb{R}^d) \) converging to \( \mu \in P_2(\mathbb{R}^d) \). Since \( \int_{\mathbb{R}^d} \langle x, y \rangle^2 d\mu(x) \) is a continuous function of \( y \in \mathbb{R}^d \), we can define

\[
y_0 = \arg\min_{y \in S^{d-1}} \int_{\mathbb{R}^d} \langle x, y \rangle^2 d\mu(x).
\]

Since

\[
\langle x, y_0 \rangle^2 \leq \|x\|^2 \|y_0\|^2 \leq \|y_0\|^2 (1 + \|x\|^2),
\]

by definition of weak convergence in \( P_2(\mathbb{R}^d) \),

\[
\int_{\mathbb{R}^d} \langle x, y_0 \rangle^2 d\mu_n(x) \to \int_{\mathbb{R}^d} \langle x, y_0 \rangle^2 d\mu(x).
\]

Since for all \( n \), the values of \( \int_{\mathbb{R}^d} \langle x, y_0 \rangle^2 d\mu_n(x) \) are bounded above and below by \( B \) and \( A \), respectively, \( \mu \) is an element of \( \text{PF}(A, B, \mathbb{R}^d) \).

Taking \( A = B \), we also have the closedness of \( \text{PF}(A, A, \mathbb{R}^d) \), the set of tight probabilistic frames with frame bound \( A \). \( \square \)

The set of probabilistic frames itself is not closed, since one can construct a sequence of probabilistic frames whose lower frame bounds converge to zero: for example, a sequence of zero-mean, Gaussian measures with covariances \( \frac{1}{n} I, n \in \mathbb{N} \).

1.1. Tightness and Scaling in the Probabilistic Setting. Tight frames are those for which the frame bounds are equal, which is equivalent to having a frame operator which is a multiple of the identity. A class of interesting frames are the scalable frames: unit-norm frames, the lengths of which can be rescaled to turn the frame into a tight frame, we refer to [6, 7, 8] for more details on scalable frames. Not all unit-norm frames are scalable, and the question of determining which are scalable can be approached by changing the weights on discrete probabilistic frames in order to obtain tight probabilistic frames. This is a different perspective on the scalable frames problem dictated by the constraints of the probabilistic point of view. While the usual approach would be equivalent to scaling the magnitudes of the vectors in the support of a probabilistic frame with equal weights, in this approach, scalability is achieved by changing the relative weights given to the delta masses supported on each of the vectors in the support of the discrete probabilistic frame, without changing the lengths of the vectors themselves. Our condition for probabilistic scalability depends on the probabilistic frame potential explained below. A similar condition could be used for scalability in the finite frame case; then, the constraint on our weights (that they are nonnegative and sum to unity)
would translate to a requirement that the squares of the new lengths sum to a constant.

To identify tight probabilistic frames, we consider the probabilistic frame potential:

**Definition 2.** Given a probabilistic frame $\mu$, the probabilistic frame potential for $\mu$ is given by

$$(1) \quad PFP(\mu) = \iint_{\mathbb{R}^d \times \mathbb{R}^d} \langle x, y \rangle^2 d\mu(x) d\mu(y)$$

As a special case, we define the frame potential for a finite frame, $\Phi = \{\varphi_i\}_{i=1}^N \subset \mathbb{R}^d$, by

$$(2) \quad FP(\Phi) = \sum_{i,j=1}^N \langle \varphi_i, \varphi_j \rangle^2 = N^2 PFP(\mu_\Phi)$$

The frame potential is a well-studied object. In their celebrated paper on finite unit-norm tight frames (FUNTFs), Benedetto and Fickus establish that, among all unit-norm frames, FUNTFs are the minimizers of $\Phi = \{\varphi_i\}_{i=1}^N \subset \mathbb{R}^d$. Because FUNTFs (and tight frames in general) have a multitude of uses in pure mathematics, statistics, and coding theory, this consequently made the frame potential a very useful quantity. The frame potential and related potentials are also studied in the context of spherical t-designs. For probabilistic frames, there exists a similar result on the minimization of the probabilistic frame potential. A version of it was first proven in [4, Theorem 4.2] for frames supported on the unit sphere, which we prove here in slightly greater generality.

**Lemma 3.** Let $\mu$ be a measure in $P_2(\mathbb{R}^d)$. The following bound holds for the probabilistic frame potential: $PFP(\mu) \geq \frac{M_2^2(\mu)}{d}$.

**Proof.** Note that, writing $m_{i,j}(\mu) = \int_{\mathbb{R}^d} x_i x_j d\mu(x)$, we have:

$$PFP(\mu) = \iint_{\mathbb{R}^d \times \mathbb{R}^d} \langle x, y \rangle^2 d\mu(x) d\mu(y) = \iint_{\mathbb{R}^d \times \mathbb{R}^d} \sum_{i=1}^d \sum_{j=1}^d x_i y_i x_j y_j d\mu(x) d\mu(y) = \sum_{i=1}^d \sum_{j=1}^d m_{i,j}^2(\mu)$$
And by Hölder,

$$M_2^2(\mu) = \sum_{i=1}^{d} m_{i,i}(\mu)$$

$$\leq \left( \sum_{i=1}^{d} m_{i,i}^2(\mu) \right)^{\frac{1}{2}} \left( \sum_{i=1}^{d} 1 \right)^{\frac{1}{2}}$$

$$\leq \sqrt{d} \left( \sum_{i=1}^{d} \sum_{j=1}^{d} m_{i,j}(\mu) \right)^{\frac{1}{2}}$$

from which the result follows. \(\square\)

**Remark 4.** Clearly, minimizers exist. In particular, if \(\mu\) is a tight probabilistic frame, then equality holds in the above claim, since the frame operator is \(S_\mu = \frac{M_2^2(\mu)}{d} I\) and

$$PFP(\mu) = \iint_{\mathbb{R}^d \times \mathbb{R}^d} \langle x, y \rangle^2 d\mu(x) d\mu(y)$$

$$= \int_{\mathbb{R}^d} \langle S_\mu y, y \rangle d\mu(y)$$

$$= \frac{M_2^4(\mu)}{d}$$

**Theorem 5.** A probabilistic frame \(\mu\) with \(M_2(\mu) = 1\) is tight if and only if it is a minimizer among \(\{ \nu \in P_2(\mathbb{R}^d) : M_2(\nu) = 1 \}\) of the probabilistic frame potential.

**Proof.** The necessity is clear from Remark 4. For the sufficiency, we consider a measure \(\mu\) in \(P_2(\mathbb{R}^d)\) which minimizes the probabilistic frame potential among \(\{ \nu \in P_2(\mathbb{R}^d) : M_2(\nu) = 1 \}\). Given any \(\nu \in P_2(\mathbb{R}^d)\) with \(M_2(\nu) = 1\), and \(\lambda \in [0, 1]\), let \(\mu_\lambda := \lambda \mu + (1 - \lambda) \nu\). That is, given a test function \(f(x)\) with at most quadratic growth,

$$\int_{\mathbb{R}^d} f(x) d\mu_\lambda(x) = \lambda \int_{\mathbb{R}^d} f(x) d\mu(x) + (1 - \lambda) \int_{\mathbb{R}^d} f(x) d\nu(x).$$
Then

\[
M_2^2(\mu_\lambda) = \int_{\mathbb{R}^d} \|x\|^2d\mu_\lambda(x)
\]
\[
= (\lambda) \int_{\mathbb{R}^d} \|x\|^2d\mu(x) + (1 - \lambda) \int_{\mathbb{R}^d} \|x\|^2d\nu(x)
\]
\[
= \lambda M_2^2(\mu) + (1 - \lambda)M_2^2(\nu)
\]
\[
= 1
\]

Therefore, since it follows that \( PFP(\mu) \leq PFP(\mu_\lambda) \forall \lambda \in [0, 1] \), we obtain:

\[
0 \leq PFP(\mu_\lambda) - PFP(\mu)
\]
\[
= \int_{\mathbb{R}^d \times \mathbb{R}^d} \langle x, y \rangle^2d\mu_\lambda(x)d\mu_\lambda(y) - \int_{\mathbb{R}^d \times \mathbb{R}^d} \langle x, y \rangle^2d\mu(x)d\mu(y)
\]
\[
= (\lambda^2 - 1)PFP(\mu) + (1 - \lambda)^2PFP(\nu)
\]
\[
+ 2\lambda(1 - \lambda) \int_{\mathbb{R}^d \times \mathbb{R}^d} \langle x, y \rangle^2d\mu(x)d\nu(y)
\]
\[
= (\lambda - 1) ((\lambda + 1)PFP(\mu) - (1 - \lambda)PFP(\nu))
\]
\[
- 2\lambda(\lambda - 1) \int_{\mathbb{R}^d} \langle y, S_\mu y \rangle d\nu(y)
\]
\[
\leq (\lambda - 1) \left( \frac{(\lambda + 1)}{d} - \frac{(1 - \lambda)}{d} - 2\lambda \int_{\mathbb{R}^d} \langle y, S_\mu y \rangle d\nu(y) \right)
\]
\[
= (\lambda - 1) \left( \frac{2\lambda}{d} - 2\lambda \int_{\mathbb{R}^d} \sum_{k=1}^{d} \lambda_k \langle y, v_k \rangle^2 d\nu(y) \right)
\]

where the second inequality comes from the fact that \( PFP(\nu) \geq \frac{M_2^2(\nu)}{d} = \frac{1}{d} \) and \( PFP(\mu) = \frac{1}{d} \), and in the last equality, the values \( \{\lambda_k\}_{k=1}^{d} \) are the eigenvalues of the frame operator \( S_\mu \), and the \( \{v_k\}_{k=1}^{d} \) are the corresponding orthonormal set of eigenvectors. It follows that

\[
\int_{\mathbb{R}^d} \sum_{k=1}^{d} \lambda_k \langle y, v_k \rangle^2 d\nu(y) \geq \frac{1}{d}
\]

Let \( \lambda_1 \) denote the smallest eigenvalue of \( \mu \), and \( v_1 \) the corresponding eigenvector of \( S_\mu \). Since \( \nu \) was chosen arbitrarily in \( \{\nu \in P_2(\mathbb{R}^d) : \)
$M_2(\nu) = 1$, it follows that for any $\epsilon > 0$, one can choose $d\nu = (1 - \epsilon)\delta_{v_1} + \frac{\epsilon}{d-1} \sum_{k=2}^d \delta_{v_k}$. Then

\[
\frac{1}{d} \leq \int_{\mathbb{R}^d} \sum_{k=1}^d \lambda_k \langle y, v_k \rangle^2 d\nu(y) = (1 - \epsilon)\lambda_1 \| v_1 \|^2 + \frac{\epsilon}{d-1} \sum_{k=2}^d \lambda_k \| v_k \|^2 = (1 - \epsilon)\lambda_1 + \frac{\epsilon}{d-1} \sum_{k=2}^d \lambda_k
\]

and as $\epsilon \to 0$, we see that, in fact, $\lambda_1 \geq \frac{1}{d}$. Since $\lambda_1 \leq \frac{M_2^2(\eta)}{d}$ for any probabilistic frame $\eta$, with equality if and only if $\eta$ is tight, it follows that our minimizer of the probabilistic frame potential, $\mu$, is tight. □

In order to solve the scaling problem in the probabilistic setting, we therefore seek weighting schemes that turn probabilistic frames supported on the unit sphere into minimizers of the probabilistic frame potential.

Consider a finite frame $\{\varphi_i\}_{i=1}^N \in \mathbb{R}^d$ with $\|\varphi_i\| = 1 \ \forall i$. Let $\mu_0 = \sum_{i=1}^N \delta_{\varphi_i}$ and $\mu_A = \sum_{i=1}^N a_i \delta_{\varphi_i}$ with $\sum_{i=1}^N a_i = 1$, $a_i \geq 0$. In this case, the probabilistic frame potential functional is given by:

\[
PFP(\mu_A) = \sum_{i,j=1}^N a_i a_j \langle \varphi_i, \varphi_j \rangle^2 \geq \frac{M_2^2(\mu_A)}{d} = \frac{1}{d},
\]

with equality if and only if $\mu_A$ is tight. Defining the matrix $Q := \begin{bmatrix} \|\varphi_i\|, \|\varphi_j\| \end{bmatrix}$, we see that

\[
PFP(\mu_A) = a^T Q a, \quad \text{where} \quad a^T = \begin{bmatrix} a_1 & \cdots & a_N \end{bmatrix}.
\]

$Q$ is positive semidefinite since it is the Hadamard product of the (positive semidefinite) Grammian matrix with itself. Thus, letting $m = \text{rank}(Q)$, we can write

\[
Q = \sum_{i=1}^m \lambda_i v_i v_i^T,
\]

where $\lambda_1 \geq \cdots \geq \lambda_m > 0$ are the nonzero eigenvalues of $Q$, and $\{v_i\}_{i=1}^N \subset \mathbb{R}^N$ are orthonormal eigenvectors. Further, any vector $a \in \mathbb{R}^N$ as can be written as $a = \sum_{i=1}^N c_i v_i$ for some constants $c_i$. 
If \( a \) leads to a minimizer of PFP, then \( a^\top Q a = \sum_{i=1}^m c_i^2 \lambda_i = \frac{1}{d} \). Observe that for each \( k \in \{1, \ldots, N\} \), the diagonal of \( Q_{k,k} \) of \( Q \) satisfies \( Q_{k,k} = \| \varphi_k \|^4 = \sum_{i=1}^m \lambda_i(v_k^i)^2 = 1 \).

Thus, letting \( s_i = \sum_{k=1}^N v_k^i \), the problem of finding \( \{a_1, \ldots, a_N\} \) for which \( \mu_A \) minimizes PFP can be reduced to solving under the following constraints:

\[
\begin{align*}
\sum_{i=1}^m \lambda_i c_i^2 & = \frac{1}{d} \\
\sum_{i=1}^N c_i v_i^k & \geq 0 \\
\sum_{i=1}^m c_i s_i & = 1 - \sum_{i=m+1}^N c_i s_i
\end{align*}
\]

(1): quadratic form  
(2): nonnegativity of a  
(3): sum of entries of a

where the sums in (1) and (3) are divided up at the index corresponding to the rank of \( Q \).

Constraints (1) and (3) make this a problem of finding the intersection of a hyperplane \( H \) and an ellipsoid \( E \) in \( \mathbb{R}^m \), where the variable is the vector \( c = [c_1 \ldots c_N] \). In particular, any intersection point \( y \) should lie between two parallel hyperplanes tangent to the ellipsoid. In particular, the coordinates of \( y \) should be bounded in magnitude by the magnitudes of the coordinates of the intersection points of the hyperplanes with the coordinate axes. That is, if \( z \) is the intersection of a tangent plane with the first coordinate axis, then \( |y_1| < |z_1| \). Given a point \( u \) on \( E \), the equation of its tangent plane is

\[
(3) \quad 2 \begin{bmatrix} \lambda_1 u_1 \\ \ldots \\ \lambda_d u_d \end{bmatrix} \cdot (x - u) = 0
\]

(4) \quad \sum_{i=1}^m \lambda_i u_i x_i = \sum_{i=1}^m \lambda_i u_i^2

(5) \quad \sum_{i=1}^m \lambda_i u_i x_i = \frac{1}{d}

Thus, the \( i \)-th intercept of the tangent plane, obtained by setting \( x_j = 0 \) for all \( j \neq i \), is \( x_i = \frac{1}{d \lambda_i u_i} \). Conversely, if we have the coordinates of the intercepts of a tangent plane, we can obtain the point of tangency via \( u_i = \frac{1}{d \lambda_i x_i} \).

The equation of the hyperplane \( H_1 \) whose points satisfy constraint (3) can be written as \( s \cdot (x - t) = 0 \), with \( t = \sum_{i>m} s_i c_i - 1 \), and if it
is parallel to a tangent plane $H_2$ to $E$, then there is some $k \neq 0$ and some $u \in E$ such that $s = k \begin{bmatrix} \lambda_1 u_1 \\ \vdots \\ \lambda_m u_m \end{bmatrix}$. The point of tangency of $H_2$ is $u$, with $u_i = \frac{s_i}{k\lambda_i}$, satisfying:

$$\sum_{i=1}^{m} \lambda_i \left( \frac{s_i}{k\lambda_i} \right)^2 = \frac{1}{d}$$

and the intercepts of the $H_2$ are

$$x_i = \frac{1}{d\lambda_i u_i} = \frac{k}{ds_i}.$$  

The $i$-th intercept of $H_1$ is $x_i = -\frac{t}{s_i}$, so that from (7) and (6), we see that we must require for each $i \in \{1, \cdots, m\}$:

$$\left| -\frac{t}{s_i} \right| < \left| \frac{k}{ds_i} \right|$$

$$\left| 1 - \sum_{i \geq m} s_i c_i \right| < \frac{|k|}{d}$$

$$\left| 1 - \sum_{i \geq m} s_i c_i \right| < \sqrt{\frac{1}{d} \sum_{i=1}^{m} \frac{s_i^2}{\lambda_i}}$$

We have thus proven:

**Lemma 6.** If $a = \sum_{i=1}^{N} c_i v_i$ with $\{c_i\}_{i=1}^{N}$ and $\{v_i\}_{i=1}^{N}$ satisfying (8), then $a$ satisfies constraints (1) and (3).

We use this result to prove the following, where we denote by $Q^\dagger$ the Moore-Penrose pseudoinverse of $Q$.

**Theorem 7.** Given $\{\varphi_i\}_{i=1}^{N} \subset S^{d-1}$, let $Q \in \mathbb{R}^{N \times N}$ be the matrix defined by $Q_{i,j} = \langle \varphi_i, \varphi_j \rangle$. Then if $\sum_{i,j=1}^{N} Q_{i,j} = d$, there exists $\{a_i\}_{i=1}^{N}$ with $a_i \geq 0$, $\sum_{i=1}^{N} a_i = 1$ such that $\mu := \sum_{i=1}^{N} a_i \delta_{\varphi_i}$ is a tight probabilistic frame.

**Proof.** Again, $Q$ is symmetric, positive semi-definite. Letting $\text{rank}(Q) = m > 0$, we have $m$ positive eigenvalues $\{\lambda_i\}_{i=1}^{m}$ and an orthonormal basis of eigenvectors $\{v_i\}_{i=1}^{N}$ and $Q = \sum_{i=1}^{m} \lambda_i v_i v_i^\top$. Given $r \in \left[ \frac{1}{N^2}, \frac{1}{N} \right]$, choose a
probability vector $a \in \mathbb{R}^N$ (i.e., $a_i \geq 0$, $i \in \{1, \cdots, N\}$, $\sum_{i=1}^{N} a_i = 1$) as in Lemma 6 with $\|a\|^2 = r$.

We note that the assumption above on the entries of Moore-Penrose pseudoinverse of $Q$ is equivalent to

$$z^T Q^T z \geq \frac{d}{N^2},$$

where $z := [\frac{1}{N} \ldots \frac{1}{N}]^T \in \mathbb{R}^N$.

Now suppose $\frac{d}{N^2} \leq z^T Q^T z = \sum_{i=1}^{m} \frac{1}{\lambda_i} \langle v_i, z \rangle^2$, so that

$$\frac{N}{d} \sum_{i=1}^{m} \frac{1}{\lambda_i} \langle v_i, z \rangle^2 \geq \frac{1}{N}.$$

Then the probability vector $a$ chosen above satisfies

$$\|a\|^2 \leq \frac{N}{d} \sum_{i=1}^{m} \frac{1}{\lambda_i} \langle v_i, z \rangle^2$$

First, $\|a\|^2 = \sum_{i=1}^{N} \langle v_i, a \rangle^2 \geq \sum_{i=1}^{m} \langle v_i, a \rangle^2$ implies that

$$\sum_{i=1}^{m} \langle v_i, a \rangle^2 \leq \frac{N}{d} \sum_{i=1}^{m} \frac{1}{\lambda_i} \langle v_i, z \rangle^2$$

Second, $\|z\|^2 = \frac{1}{N}$, so that

$$\sum_{i=1}^{m} \langle v_i, z \rangle^2 \leq \frac{1}{N}$$

Thus, by (9) and (10) and the CBS inequality, we have:

$$N \sum_{i=1}^{m} \langle v_i, z \rangle^2 \sum_{i=1}^{m} \langle v_i, a \rangle^2 \leq \frac{N}{d} \sum_{i=1}^{m} \frac{1}{\lambda_i} \langle v_i, z \rangle^2$$

$$\langle N \sum_{i=1}^{m} \langle v_i, z \rangle \langle v_i, a \rangle \rangle^2 \leq \frac{N^2}{d} \sum_{i=1}^{m} \frac{1}{\lambda_i} \langle v_i, z \rangle^2$$
Then, noting that $1 - N\langle z, a \rangle = 0$, and recalling that $N\langle z, a \rangle = \sum_{i=1}^{N} \langle z, v_i \rangle \langle v_i, a \rangle$, we obtain

$$(1 - N \sum_{i>m} \langle v_i, z \rangle \langle v_i, a \rangle)^2 = (1 - N\langle z, a \rangle)$$

$$+ N \sum_{i=1}^{m} \langle v_i, z \rangle \langle v_i, a \rangle)^2$$

$$\leq N^2 \frac{1}{d} \sum_{i=1}^{m} \frac{1}{\lambda_i} \langle v_i, z \rangle^2$$

But a quick calculation shows that this last inequality is equivalent to

$$(1 - \sum_{i>m} s_i \langle v_i, a \rangle)^2 \leq \frac{1}{d} \sum_{i=1}^{m} \frac{s_i^2}{\lambda_i},$$

where, as in Lemma 6, $s_k = \sum_{i=1}^{N} v_i^k$. Thus, by that lemma, we have that $a$, in addition to satisfying constraint (2), satisfies also constraints (1) and (3). $\square$

**Corollary 8.** Let $\lambda = \lambda_{\text{max}}(Q)$. If $\lambda^2 d \leq \sum_{i,j}^{N} Q_{i,j}$, then there exists a $\mu$ such that $\mu := \sum_{i=1}^{N} a_i \delta_{\varphi_i}$ is a tight probabilistic frame.

2. **Conclusion**

At first glance, it appears that the approach from the probabilistic context, by adding the constraint on the scaling that it be a probability vector, would make the scaling problem less tractable. However, it is this very added structure which enables the result above.

**Acknowledgment**

Clare Wickman Lau would like to thank the Norbert Wiener Center for Harmonic Analysis and Applications for its support during this research. Kasso Okoudjou was partially supported by a RASA from the Graduate School of UMCP and by a grant from the Simons Foundation (#319197 to Kasso Okoudjou).
REFERENCES

[1] G. Kutyniok and P. G. Casazza, Finite Frames. New York: Springer, 2011.
[2] M. Ehler and K. A. Okoudjou, Finite Frames, Ed. G. Kutyniok and P. G. Casazza. New York: Springer, 2011, ch. Probabilistic frames: An overview.
[3] ——, “Minimization of the probabilistic p-frame potential,” Journal of Statistical Planning and Inference, vol. 142, no. 3, pp. 645–659, 2012.
[4] M. Ehler, “Random tight frames,” Journal of Fourier Analysis and Applications, vol. 18, no. 1, pp. 1–20, 2012.
[5] R. Vershynin, “How close is the sample covariance matrix to the actual covariance matrix?” Journal of Theoretical Probability, vol. 25, no. 3, pp. 655–686, 2012.
[6] G. Kutyniok, K. A. Okoudjou, F. Philipp, and K. Tuley, “Scalable frames,” Linear Algebra Appl., pp. 2225–2238, 2013.
[7] G. Kutyniok, K. A. Okoudjou, and F. Philipp, “Scalable frames and convex geometry,” Contemp. Math., pp. 19–32, 2014.
[8] J. Cahill and X. Chen, “A note on scalable frames,” in Proceedings SampTA, Bremen, 2013, pp. 93–96.
[9] J. J. Benedetto and M. Fickus, “Finite normalized tight frames,” Advances in Computational Mathematics, vol. 18, no. 2-4, pp. 357–385, 2003.

JOHNS HOPKINS UNIVERSITY, APPLIED PHYSICS LABORATORY LAUREL, MD 20723
E-mail address: clare.wickman@jhuapl.edu

UNIVERSITY OF MARYLAND DEPARTMENT OF MATHEMATICS COLLEGE PARK, MD 20742
E-mail address: kasso@math.umd.edu