

On Optimal Frame Conditioners

Chae A. Clark

Department of Mathematics
University of Maryland, College Park
Email: cclark18@math.umd.edu

Kasso A. Okoudjou

Department of Mathematics
University of Maryland, College Park
Email: kasso@math.umd.edu

Abstract—A (unit norm) frame is scalable if its vectors can be rescaled so as to result into a tight frame. Tight frames can be considered optimally conditioned because the condition number of their frame operators is unity. In this paper we reformulate the scalability problem as a convex optimization question. In particular, we present examples of various formulations of the problem along with numerical results obtained by using our methods on randomly generated frames.

I. FRAMES AND SCALABLE FRAMES

A. Introduction

A finite frame for \mathbb{R}^N is a set $\Phi = \{\varphi_k\}_{k=1}^M \subset \mathbb{R}^N$ such that there exist positive constants $0 < A \leq B < \infty$ (referred to as the frame bounds) for which

$$A\|x\|_2^2 \leq \sum_{k=1}^M |\langle x, \varphi_k \rangle|^2 \leq B\|x\|_2^2$$

for all $x \in \mathbb{R}^N$. Given a frame $\Phi = \{\varphi_k\}_{k=1}^M \subset \mathbb{R}^N$, we denote again by Φ the $N \times M$ matrix whose k^{th} column is the vector φ_k . The matrix Φ is the synthesis operator associated to the frame Φ , and its transpose Φ^T is the analysis operator of Φ . The frame operator is then defined as $S = \Phi\Phi^T$. When $A = B$ the frame is called tight, in which case the frame operator is $S = AI$ where I denotes the identity matrix of the necessary size.

B. Scalable Frames

Scalable frames were introduced in [12], [11] as a method to convert a non tight frame into a tight one. More precisely:

Definition 1. Let $M \geq N$ be given. A frame $\Phi = \{\varphi_k\}_{k=1}^M \subset \mathbb{R}^N$ is called scalable if there exist a subset $\Phi_J = \{\varphi_k\}_{k \in J} \subset \mathbb{R}^N$ with $J \subseteq \{1, 2, \dots, M\}$, and non-negative scalars $\{x_k\}_{k \in J}$ such that the system $\tilde{\Phi}_J = \{x_k \varphi_k\}_{k \in J}$ is a tight frame for \mathbb{R}^N .

This definition slightly differs from the one presented in [11, Definition 2.1] as we shall not distinguish between the scaling of frame elements, and their removal from the frame (strictly and m -scalable). Let $\Phi = \{\varphi_k\}_{k=1}^M \subset \mathbb{R}^N$ be a frame. Then the analysis operator of the scaled frame $\{x_k \varphi_k\}_{k=1}^M$ is given by $X\Phi^T$, where X is the diagonal matrix with the values x_k on its diagonal. Hence, the frame Φ is scalable if and only

if there exists a diagonal matrix $X = \text{diag}(x_k)$, with $x_k \geq 0$ such that

$$\tilde{S} = \Phi X^T X \Phi^T = \Phi X^2 \Phi^T = AI. \quad (\text{I.1})$$

for some constant $A > 0$.

One can convert (I.1) into a linear system of equations in M unknowns: x_k^2 . To write out this linear system we need the following function: $F : \mathbb{R}^N \rightarrow \mathbb{R}^d$ given by

$$F(x) = [F_0(x), F_1(x), \dots, F_{N-1}(x)]^T,$$

$$F_0(x) = \begin{bmatrix} x_1^2 - x_2^2 \\ x_1^2 - x_3^2 \\ \vdots \\ x_1^2 - x_N^2 \end{bmatrix}, F_k(x) = \begin{bmatrix} x_k x_{k+1} \\ x_k x_{k+2} \\ \vdots \\ x_k x_N \end{bmatrix}$$

and $F_0(x) \in \mathbb{R}^{N-1}$, $F_k(x) \in \mathbb{R}^{N-k}$, $k = 1, 2, \dots, N-1$, where $d := \frac{(N-1)(N+2)}{2}$. Let $F(\Phi)$ be the $d \times M$ matrix given by

$$F(\Phi) = (F(\varphi_1) \ F(\varphi_2) \ \dots \ F(\varphi_M)).$$

In this setting we have the following solution to the scalability problem:

Proposition 2. [11, Proposition 3.7] A frame $\Phi = \{\varphi_k\}_{k=1}^M \subset \mathbb{R}^N$ is scalable if and only if there exists a non-negative $u \in \ker F(\Phi) \setminus \{0\}$.

C. Mathematical Programming and Duality

Our main goal is to find a non-negative nontrivial vector in the null space of $F(\Phi)$ using some optimization methods. For this reason we recall some notions from duality theory in mathematical programming. For a more robust treatment of duality theory applied to linear programs, we refer to the standard texts by Boyd, Vandenberghe [3] and Bertsimas, Tsitsiklis [2]. Recall that the Primal and Dual linear mathematical programming problems are defined, respectively, as follows:

$$\begin{aligned} &\text{minimize: } c^T x \\ &\text{subject to: } Ax = b \\ & \quad \quad \quad x \geq 0. \end{aligned}$$

$$\begin{aligned} &\text{maximize: } b^T y \\ &\text{subject to: } A^T y \leq c \\ & \quad \quad \quad y \in \mathbb{R}^N. \end{aligned}$$

Theorem 3 (Strong Duality). If either the primal or dual problem has a finite optimal value, then so does the other. The optimal values coincide, and optimal solutions to both the primal and dual problems exist.

Theorem 4 (Complimentary Slackness). Let x^* and y^* be feasible solutions to the primal and dual problems respectively. Let A be an N by M matrix, where A_j denotes the j th column and a_i denotes the i th row of A . Then x^* and y^* are optimal solutions to their respective problems if and only if

$$y_i(a_i \cdot x - b_i) = 0 \text{ for all } i = 1, \dots, N,$$

and

$$x_j(c_j - y^T A_j) = 0 \text{ for all } j = 1, \dots, M.$$

II. REFORMULATION OF THE SCALABILITY PROBLEM AS AN OPTIMIZATION PROBLEM

This section establishes the equivalence of generating a scaling matrix X and solving an optimization problem of a generic convex objective function. More specifically, we shall phrase the scalability problem as a linear and convex programming problem.

First consider the sets \mathcal{S}_1 and \mathcal{S}_2 given by

$$\mathcal{S}_1 := \{u \in \mathbb{R}^M \mid F(\Phi)u = 0, u \geq 0, u \neq 0\},$$

and

$$\mathcal{S}_2 := \{v \in \mathbb{R}^M \mid F(\Phi)v = 0, v \geq 0, \|v\|_1 = 1\}.$$

\mathcal{S}_1 is a subset of the null space of $F(\Phi)$, and each $u \in \mathcal{S}_1$ is associated a scaling matrix X_u , defined as

$$X_u := (X_{ij})_u = \begin{cases} \sqrt{u_i} & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

$\mathcal{S}_2 \subset \mathcal{S}_1 \cap B_{\ell^1}$ where B_{ℓ^1} is the unit ball under the ℓ^1 norm.

We observe that a frame $\Phi = \{\varphi_k\}_{k=1}^M \subset \mathbb{R}^N$ is scalable if and only if there exists a scaling matrix X_u with $u \in \mathcal{S}_1$. Consequently, one can associate to X_u a scaling matrix X_v with $v \in \mathcal{S}_2$. The normalized set \mathcal{S}_2 ensures that the constraints in the optimization problems to be presented are convex.

Theorem 5. Let $\Phi = \{\varphi_k\}_{k=1}^M \subset \mathbb{R}^N$ be a frame, and let $f : \mathbb{R}^M \rightarrow \mathbb{R}$ be a convex function. Then the program

$$\begin{aligned} (\mathcal{P}) \text{ minimize: } & f(u) & (\text{II.1}) \\ \text{subject to: } & F(\Phi)u = 0 \\ & \|u\|_1 = 1 \\ & u \geq 0 \end{aligned}$$

has a solution if and only if the frame Φ is scalable.

Proof. Any feasible solution u^* of \mathcal{P} is contained in the set \mathcal{S}_2 , which itself is contained in \mathcal{S}_1 , and thus corresponds to a scaling matrix X_{u^*} .

Conversely, any $u \in \mathcal{S}_1$ can be mapped to a $v \in \mathcal{S}_2$ by appropriate scaling factor. This provides an initial feasible solution to \mathcal{P} , and as f is convex and the constraints are convex and bounded, there must exist a minimizer of \mathcal{P} . \square

Theorem 5 is very general in that the convex objective function f can be chosen so as the resulting frame has some desirable properties. We now consider certain interesting examples of objective functions f . These examples can be related to the sparsity (or lack thereof) of the desired solution. Using a linear objective function promotes sparsity, while barrier objectives promote dense solutions (small number of zero elements in u).

A. Linear Program Formulation

Assume that the objective function in (II.1) is given by $f(u) := a^T u$ for some coefficient vector $a \in \mathbb{R}^M \setminus \{0\}$. Our program \mathcal{P} now becomes

$$\begin{aligned} (\mathcal{P}_1) \text{ minimize: } & a^T u & (\text{II.2}) \\ \text{subject to: } & F(\Phi)u = 0 \\ & \|u\|_1 = 1 \\ & u \geq 0. \end{aligned}$$

Choosing the coefficients a independently of the variables u , results in a linear program. For example, the choice $a_i = 1$ for all i result in a program to minimize the ℓ^1 norm of u . Another, more useful choice of coefficients is $a_i = \frac{1}{\|F(\varphi_i)\|_2}$. Under this regime, a higher weight is given to the frame elements with smaller norm (which further encourages sparsity).

One of the advantages of linear programs is that they admit a strong dual formulation. To the primal problem \mathcal{P}_1 corresponds the following dual problem \mathcal{P}_2 .

Proposition 6. Let $\Phi = \{\varphi_k\}_{k=1}^M \subset \mathbb{R}^N$ be a frame. The program

$$\begin{aligned} (\mathcal{P}_2) \text{ maximize: } & w \\ \text{subject to: } & [F(\Phi)^T \mathbb{1}] \begin{bmatrix} v \\ w \end{bmatrix} \leq a \\ & w \in \mathbb{R}, v \in \mathbb{R}^d \end{aligned}$$

is the strong dual of \mathcal{P}_1 .

Proof. This result follows exactly from the construction of dual formulations for linear programs. The primal problem can be formulated as follows:

$$\begin{aligned} \text{minimize: } & \sum_{i=1}^M a_i u_i \\ \text{subject to: } & F(\Phi)u = 0 \\ & \sum_{i=1}^M u_i = 1 \\ & u \geq 0. \end{aligned}$$

The strong dual of this problem is:

$$\begin{aligned} \text{maximize: } & w \\ \text{subject to: } & [F(\Phi)^T \mathbb{1}] \begin{bmatrix} v \\ w \end{bmatrix} \leq a. \end{aligned}$$

\square

Numerical optimization schemes, in many cases, consist of a search for an initial feasible solution, and then a search for an optimal solution. In analyzing the linear program formulation \mathcal{P}_1 , we notice that we either have an optimal solution or the problem is infeasible, but there is no case when the problem is unbounded (due to the bounding constraint $\|u\|_1 = 1$).

The dual problem has the property that it either has an optimal solution, or is unbounded (from duality). Consequently, for any frame Φ , $w = \min\{a\}$ and $v = 0$ is always a feasible solution to the dual problem. This removes the requirement that an initial solution be found [2].

B. Barrier Formulations

A sparse solution to the linear program produces a frame in which the frame elements corresponding to the zero coefficients are removed. In contrast, one may wish to have a full solution, that is, one may want to retain all of, or most of, the frame vectors. To enforce this property, we use two types of barrier objective functions.

Proposition 7. Let $\Phi = \{\varphi_k\}_{k=1}^M \subset \mathbb{R}^N$ be a frame, and define $0 \leq \epsilon \ll 1$. If the problem

$$\begin{aligned} (\mathcal{P}_3) \text{ maximize: } & \sum_{i=1}^M \ln(u_i + \epsilon) & (\text{II.3}) \\ \text{subject to: } & F(\Phi)u = 0 \\ & \|u\|_1 = 1 \\ & u \geq 0. \end{aligned}$$

has a feasible solution u^* with a finite objective function value, then the frame Φ is scalable, and the scaling matrix X is a diagonal operator where the elements are the square-roots of the feasible solution u^* . Moreover, for $\epsilon = 0$, if a solution u^* exists, all elements of u^* are strictly positive.

Proof. Assume u^* is a feasible solution to (II.3) with $0 < \epsilon \ll 1$ and the objective function finite. Then from Theorem 5, we have that the frame Φ is scalable. Now assume $\epsilon = 0$. If one of the variables u_i were zero, then the objective function would have a value of $-\infty$. Since we assume the function is finite, this cannot be the case. A negative value for u_i would result in the objective function value being undefined, this also cannot be the case due to the finite objective. Therefore, u_i must be positive for all i . \square

Proposition 8. Let $\Phi = \{\varphi_k\}_{k=1}^M \subset \mathbb{R}^N$ be a frame. If the problem

$$\begin{aligned} (\mathcal{P}_4) \text{ maximize: } & \min_{i=1, \dots, M} \{u_i\} & (\text{II.4}) \\ \text{subject to: } & F(\Phi)u = 0 \\ & \|u\|_1 = 1 \\ & u \geq 0. \end{aligned}$$

has a feasible solution u^* with a finite objective function value, then the frame Φ is scalable, and the scaling matrix X is a diagonal operator where the elements are the square-roots of the feasible solution u^* . Moreover, a solution exists

with positive elements if and only if the solution produced by solving this problem has positive elements.

Proof. To show this, we shall rewrite this problem as a linear program.

$$\begin{aligned} \text{maximize: } & t & (\text{II.5}) \\ \text{subject to: } & F(\Phi)u = 0 \\ & \sum_{i=1}^M u_i = 1 \\ & t \leq u_i \\ & t > 0, u \geq 0. \end{aligned}$$

Here, t is an auxiliary variable, taken to be the minimum element of u . This linear program can be solved to optimality. Moreover, as this problem is convex, the optimum achieved is global. If the objective function at optimality has a value of 0, then there can exist no solution with all positive coefficients. \square

III. AUGMENTED LAGRANGIAN METHODS

To efficiently solve these convex formulations, we employ the method of augmented Lagrangians. Utilizing this method is more a personal choice, as a multitude of first-order methods exist. The augmented Lagrangian is chosen due to its flexibility. Sparse solutions can be induced through thresholding, and dense solutions can be induced through choice of objective function. As an example, consider the problem (\mathcal{P}) , with the objective function $f(u) := \|u\|_{\ell^2}^2$. Rewriting norms as matrix/vector products we are interested in

$$\begin{aligned} \text{minimize: } & u^T I u \\ \text{subject to: } & F(\Phi)u = 0 \\ & \mathbb{1}^T u = 1 \\ & u \geq 0. \end{aligned}$$

Here, I is the $M \times M$ identity matrix. For notational convenience, we denote L and b to be

$$\begin{bmatrix} F(\Phi) \\ \mathbb{1}^T \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

respectively. The ℓ^2 problem is now

$$\begin{aligned} \text{minimize: } & u^T I u \\ \text{subject to: } & Lu = b \\ & u \geq 0. \end{aligned}$$

To solve this problem, the augmented Lagrangian \mathcal{L} is formed,

$$\begin{aligned} \mathcal{L} &= u^T I u + \langle \mu, Lu - b \rangle + \frac{\lambda}{2} \|Lu - b\|_2^2 \\ &= u^T I u + \mu^T Lu - \mu^T b + \frac{\lambda}{2} (u^T L^T Lu - 2u^T L^T b + b^T b). \end{aligned}$$

This function is minimized through satisfying the first-order condition, $\nabla \mathcal{L} = 0$.

The gradient of the Lagrangian with respect to u is solved through standard calculus-based methods. The Lagrangian

with respect to the dual variables μ and λ is linear, which we optimize through gradient descent. The tuning parameter η denotes the scaling of the descent direction.

$$\begin{aligned}\nabla_u \mathcal{L} &= 2Iu + L^T \mu + \lambda L^T L u - \lambda L^T b. \\ 0 &= (2I + \lambda L^T L)u + L^T \mu - \lambda L^T b. \\ (2I + \lambda L^T L)u &= \lambda L^T b - L^T \mu.\end{aligned}$$

Dividing the equation by λ , we have

$$\begin{aligned}\left(\frac{2}{\lambda}I + L^T L\right)u &= L^T b - L^T \frac{\mu}{\lambda}. \\ \left(\frac{2}{\lambda}I + L^T L\right)u &= L^T \left(b - \frac{\mu}{\lambda}\right). \\ u &= \left(\frac{2}{\lambda}I + L^T L\right)^{-1} L^T \left(b - \frac{\mu}{\lambda}\right).\end{aligned}$$

The dual variables have the following gradients,

$$\begin{aligned}\nabla_\mu \mathcal{L} &= Lu - b. \\ \nabla_\lambda \mathcal{L} &= \frac{1}{2} \langle Lu - b, Lu - b \rangle.\end{aligned}$$

And forming the gradient descent algorithm with η , results in

Algorithm 1 Gradient Descent w.r.t. μ

while not converged **do**
 $\mu_{k+1} \leftarrow \mu_k - \eta \cdot (Lu - b)$
end while

Algorithm 2 Gradient Descent w.r.t. λ

while not converged **do**
 $\lambda_{k+1} \leftarrow \lambda_k - \frac{\eta}{2} \cdot \langle Lu - b, Lu - b \rangle$
end while

Lastly, to retain the non-negativity of the solution, we project the current solution onto \mathbb{R}_+ . This is accomplished by setting any negative values in the solution to 0 (thresholding). We shall denote this $\mathcal{P}_+(\cdot)$. Forming the full augmented Lagrangian scheme, we now have the complete ℓ^2 derivation.

Algorithm 3 Full Augmented Lagrangian Scheme (ℓ^2)

while not converged **do**
 $v_{k+1} \leftarrow \left(\frac{2}{\lambda_k}I + L^T L\right)^{-1} L^T \left(b - \frac{\mu_k}{\lambda_k}\right)$
 $u_{k+1} \leftarrow \mathcal{P}_+(v_{k+1})$
 $\mu_{k+1} \leftarrow \mu_k - \eta \cdot (Lu_{k+1} - b)$
 $\lambda_{k+1} \leftarrow \lambda_k - \frac{\eta}{2} \cdot \langle Lu_{k+1} - b, Lu_{k+1} - b \rangle$
end while

IV. NUMERICAL EXAMPLES

The following numerical tests are intended to illustrate our methods by scaling frames generated from Gaussian distributions. In particular, throughout this section, we identify random frames Φ with full-rank random $N \times M$ matrices whose elements are i.i.d., drawn from a Gaussian distribution with zero mean and unit variance. These random frames are not tight, but they are not badly conditioned. This follows from the incoherence between Gaussian frame elements [8] and results on singular/eigen-values by Edelman [9]. Future work will involve specific subclasses of frames to better characterize the space.

The first set of figures are intended to give a representation of how the scaling affects frames in \mathbb{R}^2 . A number of Gaussian random frames are generated in MatLab, and a scaling process is performed by solving one of the optimization problems above (the specific program used is noted under the figures). The Gaussian frame is first normalized to be on the unit circle. The (blue\circle) vectors correspond to the original frame vectors, and the (red\triangle) vectors represent the resulting scaled frame.

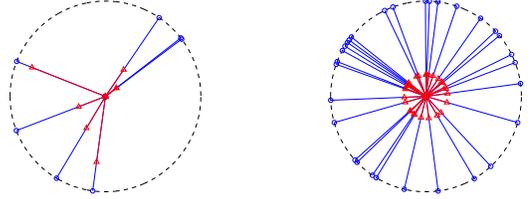


Fig. IV.1. These examples display the effect of scaling frames in \mathbb{R}^2 . The frames are sized $M = 7$ (Left) and $M = 30$ (Right), and were scaled using the Augmented Lagrangian Scheme. The left figure shows that scalings favor isolated frame elements. The right figure shows that as the frame elements fill the space, the scalings become more normalized.

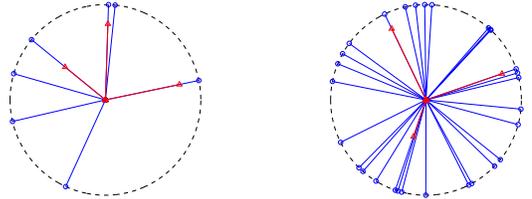


Fig. IV.2. These examples illustrate scalable frames with a small number of positive weights. The frames are sized $M = 7$ (Left) and $M = 30$ (Right), and were scaled using linear programming formulation \mathcal{P}_1 (more specifically, the Simplex algorithm). These two example show that for frames of low (Left) and high (Right) redundancy, sparse solutions are possible and seemingly unrelated to the number of frame elements.

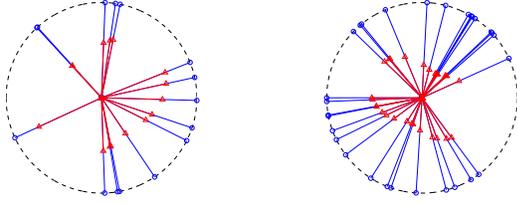


Fig. IV.3. These frames show full scaling results from the log-barrier method. The frames are sized $M = 15$ (Left) and $M = 20$ (Right), and as was mentioned in figure IV.1, the scalings favor isolated frame elements.

The next tables illustrate the sparsity that is achieved in scaling a frame. That is, using the linear program, we present the average number of non-zero frame elements retained over 100 trials. Our data seem to suggest the existence of a minimum number of frame elements required to perform a scaling, and this number seems to depend only on the dimension of the underlying space. This phenomenon should be compared to the estimates on the probability that a frame of M vectors in \mathbb{R}^N is scalable that were obtained in [7, Theorem 4.9].

Sparsity Test Results [Gaussian Frames]

$N \setminus M$	3	4	5	10	20	30	40	50
2	3	3	3	3	3	3	3	3
3	-	-	-	6.01	6	6	6	6
4	-	-	-	10.1	10.12	10.1	10	10
5	-	-	-	-	15.08	15.12	15.11	15.05

TABLE IV.1

THE AVERAGE NUMBER OF FRAME ELEMENTS RETAINED AFTER SCALING USING THE LINEAR PROGRAM FORMULATION. ENTRIES WITH A “-” IMPLY THAT THE PROPORTION OF SCALABLE FRAMES IN THE SPACE IS TOO SMALL FOR PRACTICAL TESTING.

Sparsity Test Results [Gaussian Frames]

$N \setminus M$	150	200	250	500	750	1000
10	56.06	56.02	55.72	55.57	55.66	55.6
15	-	-	123.76	123.98	123.37	123.1
20	-	-	-	217.6	218.6	219.45
25	-	-	-	-	-	338.67

TABLE IV.2

THE AVERAGE NUMBER OF FRAME ELEMENTS RETAINED AFTER SCALING USING THE LINEAR PROGRAM FORMULATION. ENTRIES WITH A “-” IMPLY THAT THE PROPORTION OF SCALABLE FRAMES IN THE SPACE IS TOO SMALL FOR PRACTICAL TESTING.

Observe that the results presented in tables IV.1 and IV.2 show an interesting trend. The average number of elements required to scale a frame appears to be $d + 1 = \frac{N(N+1)}{2}$. The linear system being solved during the Simplex method has dimensions $(d + 1) \times M$, and attempts to find a non-negative solution in \mathbb{R}^M . It seems unlikely that this solution can be found using less than $d + 1$ frame elements.

The final test presents the proportion of scalable Gaussian frames of a given size over 100 trials. For testing, the number of frame vectors is determined by the dimension of the underlying space. For a frame in \mathbb{R}^N , the number of frame elements used ranges from $N + 1$ to $4N^2$ (e.g. for $N = 2$, the number of frame elements range from $M = 3$ to $M = 16$).

A Gaussian frame is generated of the required sizes and a scaling is attempted. This is performed over a hundred trials, and the proportion of frames that were scalable was retained. For each N , a plot of the proportions across frame size M is presented. To display these plots in a single figure, the independent variable M is scaled to lie in the range (0,1).

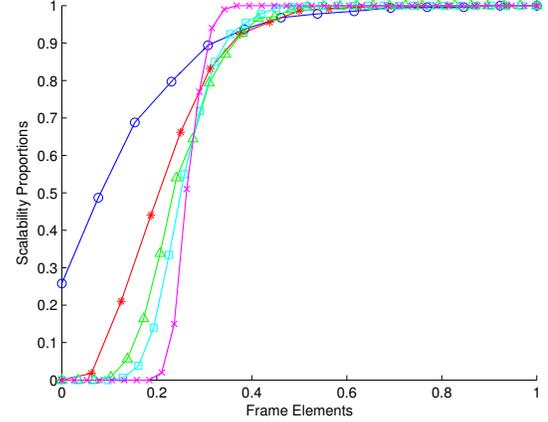


Fig. IV.4. Each graph in this figure gives the proportion of scalable frames generated after 100 trials. The sizes of the frames range from $N + 1$ to $4N^2$. To fit the graphs in a single figure, the range of each figure is scaled to be from 0 to 1. The frame dimensions are $N = 2$ (Blue\Circle), $N = 3$ (Red\Star), $N = 4$ (Green\Triangle), $N = 5$ (Cyan\Box), and $N = 10$ (Magenta\X).

V. ACKNOWLEDGMENTS

Chae Clark 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] F. Bach, R. Jenatton, J. Mairal, and G. Obozinski. Convex optimization with sparsity-inducing norms. *Optimization for Machine Learning*, pages 19–53, 2011.
- [2] D. Bertsimas and J. Tsitsiklis. Introduction to linear optimization. 1997.
- [3] S. P. Boyd and L. V. *Convex optimization*. Cambridge university press, 2004.
- [4] E. J. Candes and T. Tao. Decoding by linear programming. *Information Theory, IEEE Transactions on*, 51(12):4203–4215, 2005.
- [5] P. Casazza and G. Kutyniok. *Finite frames*. Springer, 2012.
- [6] S. S. Chen, D. Donoho, and M. Saunders. Atomic decomposition by basis pursuit. *SIAM journal on scientific computing*, 20(1):33–61, 1998.
- [7] X. Chen, G. Kutyniok, K. Okoudjou, F. Philipp, and R. Wang. Measures of scalability. *arXiv preprint arXiv:1406.2137*, 2014.
- [8] David L Donoho and Xiaoming Huo. Uncertainty principles and ideal atomic decomposition. *Information Theory, IEEE Transactions on*, 47(7):2845–2862, 2001.
- [9] Alan Edelman. Eigenvalues and condition numbers of random matrices. *SIAM Journal on Matrix Analysis and Applications*, 9(4):543–560, 1988.
- [10] G. Golub and C. V. L. *Matrix computations*, volume 3. JHU Press, 2012.
- [11] G. Kutyniok, K. Okoudjou, and F. Philipp. Scalable Frames and Convex Geometry. *Contemp. Math.* 626 (2014), 19-32.
- [12] G. Kutyniok, K. Okoudjou, F. Philipp, and E. Tuley. Scalable frames. *Linear Algebra and its Applications*, 438(5):2225 – 2238, 2013.
- [13] WR Madych. On underdetermined systems. In *Numerical Linear Algebra, Digital Signal Processing and Parallel Algorithms*, pages 541–545. Springer, 1991.