

Gradient descent of the frame potential

Peter G. Casazza ⁽¹⁾ and Matthew Fickus ⁽²⁾

(1) Department of Mathematics, University of Missouri, Columbia, MO 65211 USA.

(2) Department of Mathematics & Statistics, Air Force Institute of Technology, WPAFB, OH 45433 USA.

pete@math.missouri.edu, matthew.fickus@afit.edu

Abstract:

Unit norm tight frames provide Parseval-like decompositions of vectors in terms of possibly nonorthogonal collections of unit norm vectors. One way to prove the existence of unit norm tight frames is to characterize them as the minimizers of a particular energy functional, dubbed the frame potential. We consider this minimization problem from a numerical perspective. In particular, we discuss how by descending the gradient of the frame potential, one, under certain conditions, is guaranteed to produce a sequence of unit norm frames which converge to a unit norm tight frame at a geometric rate. This makes the gradient descent of the frame potential a viable method for numerically constructing unit norm tight frames.

1. Introduction

The *analysis* operator of some finite sequence of vectors $\{f_m\}_{m=1}^M$ in an N -dimensional Hilbert space \mathbb{H}_N is the operator $F : \mathbb{H}_N \rightarrow \mathbb{C}^M$, $(Ff)(m) := \langle f, f_m \rangle$. The corresponding *frame* operator is $F^*F : \mathbb{H}_N \rightarrow \mathbb{H}_N$,

$$F^*Ff = \sum_{m=1}^M \langle f, f_m \rangle f_m.$$

Generally speaking, *frame theory* is the study of how $\{f_m\}_{m=1}^M$ may be chosen in order to guarantee that F^*F is well-conditioned. In particular, $\{f_m\}_{m=1}^M$ is a *frame* for \mathbb{H}_N if there exists *frame bounds* $0 < A \leq B < \infty$ such that $A\mathbf{I} \leq F^*F \leq B\mathbf{I}$, and is a *tight frame* if $A = B$, that is, if $F^*F = A\mathbf{I}$.

Typically, one's choice of f_m 's is restricted according to some nonlinear, application-specific constraints. Of particular interest is the case of *unit norm tight frames*, that is, tight frames for which $\|f_m\| = 1$ for all $m = 1, \dots, M$; such frames, known to exist for any $M \geq N$, provide Parseval-like decompositions in terms of vectors of unit length, even though these vectors are possibly nonorthogonal. Despite an ever-growing list of specific constructions of such frames, little is known about the manifold structure of the set of all unit norm tight frames.

In the hunt for unit norm tight frames, the *frame potential*, specifically defined as:

$$\text{FP}(\{f_m\}_{m=1}^M) := \sum_{m,m'=1}^M |\langle f_m, f_{m'} \rangle|^2$$

for any sequence $\{f_m\}_{m=1}^M \in \mathbb{H}_N^M$, is a useful tool. Specifically, the frame potential quantifies the *total orthogonality* of a system of vectors by measuring the total potential energy stored within that system under a certain force which encourages orthogonality. Regarded as a functional over

$$\mathbb{S}_N^M = \{\{f_m\}_{m=1}^M \in \mathbb{H}_N^M : \|f_m\| = 1, m = 1, \dots, M\},$$

one may show that when $M \geq N$, the local minimizers of the frame potential are precisely the unit norm tight frames of M elements for \mathbb{H}_N . In particular, as the frame potential is continuous and \mathbb{S}_N^M is compact, one may conclude that such frames indeed exist for any $M \geq N$.

In this paper, we consider the minimization of the frame potential from a numerical perspective. In particular, in the next section, we compute the gradient of the frame potential, namely a specific direction $\{g_m\}_{m=1}^M \in \mathbb{H}_N^M$ in which to push $\{f_m\}_{m=1}^M$ so as to achieve the greatest instantaneous decrease of FP. Then, in an improvement over typical uses of gradient descent, we compute an exact step size in which to travel in this direction so as to produce a certain decrease in potential. In the third section, we estimate the size of this decrease in relation to how far the frame potential is from its minimum; under sufficient conditions, this estimate may be used to show that by descending the gradient of the frame potential, one may produce a sequence of unit norm frames which converge to a unit norm tight frame at a geometric rate.

The frame potential was introduced in [1], with its domain of optimization being later generalized in [4]. It has been used to characterize tight filter bank frames [5, 6]. The frame potential may also be used to prove the existence of tight fusion frames [3], and the local minimizers of the fusion frame potential are themselves a subject of interest [7, 9]. Further generalizations of the frame potential are considered in [2, 8].

2. The gradient of the frame potential

Our goal is to numerically minimize the frame potential over \mathbb{S}_N^M . As our domain of optimization is a product of spheres as opposed to the entire space \mathbb{H}_N^M , our approach departs from the classical theory of gradients. In particular, given $\{f_m\}_{m=1}^M \in \mathbb{S}_N^M$ and any $\{g_m\}_{m=1}^M \in \mathbb{H}_N^M$ such that $\langle f_m, g_m \rangle = 0$ for all $m = 1, \dots, M$, we shall compute the rate of change of the frame potential as each f_m

is pushed along a great circle with tangent velocity g_m . We then define the gradient of FP to be that particular $\{g_m\}_{m=1}^M$ which makes this directional derivative as large as possible. We begin with the following result, which gives the first two derivatives of the frame potential of a single parameter family of frames:

Lemma 1 (Lemma 2 of [3]). *For any set of twice-differentiable parameterized curves $\{f_m(\cdot)\}_{m=1}^M$ in \mathbb{H}_N , the first two derivatives of $\varphi(t) := \text{FP}(\{f_m(t)\}_{m=1}^M)$ are:*

$$\begin{aligned}\dot{\varphi}(t) &= 4\text{ReTr}(\dot{F}(t)F^*(t)F(t)F^*(t)), \\ \ddot{\varphi}(t) &= 4\text{ReTr}(\ddot{F}(t)F^*(t)F(t)F^*(t)) \\ &\quad + 4\|\dot{F}(t)F^*(t)\|_{\text{HS}}^2 \\ &\quad + 2\|\dot{F}^*(t)F(t) + F^*(t)\dot{F}(t)\|_{\text{HS}}^2,\end{aligned}$$

where $\dot{F}(t)$ and $\ddot{F}(t)$ are the analysis operators of $\{\dot{f}_m(t)\}_{m=1}^M$ and $\{\ddot{f}_m(t)\}_{m=1}^M$, respectively.

We now use Lemma 1 along with Taylor's theorem to asymptotically estimate the change in frame potential one obtains by perturbing a given $\{f_m\}_{m=1}^M \in \mathbb{S}_N^M$ along any choice of great circles. To be precise, letting:

$$\oplus f_m^\perp := \{g_m\}_{m=1}^M \in \mathbb{H}_N^M : \langle f_m, g_m \rangle = 0, \forall m\},$$

we have the following:

Theorem 2. *For any $\{f_m\}_{m=1}^M \in \mathbb{S}_N^M$, $\{g_m\}_{m=1}^M \in \oplus f_m^\perp$, let:*

$$f_m(t) := \cos(\|g_m\|t)f_m + (\sin(\|g_m\|t)/\|g_m\|)g_m$$

whenever $g_m \neq 0$ and let $f_m(t) := f_m$ otherwise. Then, $\{f_m(t)\}_{m=1}^M \in \mathbb{S}_N^M$ for any $t \in \mathbb{R}$, and satisfies:

$$\sum_{m=1}^M \|f_m(t) - f_m\|^2 \leq t^2 \sum_{m=1}^M \|g_m\|^2, \quad (1)$$

as well as:

$$\begin{aligned}\text{FP}(\{f_m(t)\}_{m=1}^M) &\leq \text{FP}(\{f_m\}_{m=1}^M) \\ &\quad + 4t\text{Re} \sum_{m=1}^M \langle F^*Ff_m, g_m \rangle \\ &\quad + 8Mt^2 \sum_{m=1}^M \|g_m\|^2.\end{aligned} \quad (2)$$

Proof. It is straightforward to show that $\|f_m(t)\| = 1$ for all $m = 1, \dots, M$ and all $t \in \mathbb{R}$. To show (1), note that for any m such that $g_m \neq 0$, we have:

$$\begin{aligned}\|f_m(t) - f_m\|^2 &= (\cos(\|g_m\|t) - 1)^2 + \sin^2(\|g_m\|t) \\ &= 4\sin^2(\|g_m\|t/2) \\ &\leq \|g_m\|^2 t^2.\end{aligned} \quad (3)$$

As (3) also immediately holds for any m such that $g_m = 0$, we may sum (3) over all m to conclude (1). To show (2), we apply Taylor's theorem to $\varphi(t) = \text{FP}(\{f_m(t)\}_{m=1}^M)$ at $t = 0$:

$$\varphi(t) \leq \varphi(0) + t\dot{\varphi}(0) + \frac{1}{2}t^2 \max_{s \in \mathbb{R}} |\ddot{\varphi}(s)|. \quad (4)$$

To compute the terms in (4), note that

$$\dot{f}_m(t) = -\|g_m\| \sin(\|g_m\|t)f_m + \cos(\|g_m\|t)g_m \quad (5)$$

for any m such that $g_m \neq 0$. As (5) also immediately holds when $g_m = 0$, we have $\dot{f}_m(0) = g_m$ for all m . Thus, by Lemma 1,

$$\begin{aligned}\dot{\varphi}(0) &= 4\text{ReTr}(\dot{F}(0)F^*(0)F(0)F^*(0)) \\ &= 4\text{ReTr}(\dot{F}(0)F^*FF^*) \\ &= 4\text{Re} \sum_{m=1}^M \langle \dot{F}(0)F^*FF^*e_m, e_m \rangle \\ &= 4\text{Re} \sum_{m=1}^M \langle F^*Ff_m, \dot{f}_m(0) \rangle \\ &= 4\text{Re} \sum_{m=1}^M \langle F^*Ff_m, g_m \rangle.\end{aligned} \quad (6)$$

Next, as taking the derivative of (5) yields $\ddot{f}_m(t) = -\|g_m\|^2 f_m(t)$ for any m , we have:

$$\begin{aligned}\text{Tr}(\ddot{F}(t)F^*(t)F(t)F^*(t)) &= \sum_{m=1}^M \langle \ddot{F}(t)F^*(t)F(t)F^*(t)e_m, e_m \rangle \\ &= \sum_{m=1}^M \langle F^*(t)F(t)f_m(t), \ddot{f}_m(t) \rangle \\ &= \sum_{m=1}^M \langle F^*(t)F(t)f_m(t), -\|g_m\|^2 f_m(t) \rangle \\ &= -\sum_{m=1}^M \|g_m\|^2 \|F(t)f_m(t)\|^2.\end{aligned} \quad (7)$$

In particular, combining (7) with Lemma 1 gives:

$$\begin{aligned}\ddot{\varphi}(t) &= -4 \sum_{m=1}^M \|g_m\|^2 \|F(t)f_m(t)\|^2 \\ &\quad + 4\|\dot{F}(t)F^*(t)\|_{\text{HS}}^2 \\ &\quad + 2\|\dot{F}^*(t)F(t) + F^*(t)\dot{F}(t)\|_{\text{HS}}^2.\end{aligned} \quad (8)$$

To bound (8), note that by (5),

$$\begin{aligned}\|F(t)\|_{\text{HS}}^2 &= \sum_{m=1}^M \|f_m(t)\|^2 = M, \\ \|\dot{F}(t)\|_{\text{HS}}^2 &= \sum_{m=1}^M \|\dot{f}_m(t)\|^2 = \sum_{m=1}^M \|g_m\|^2,\end{aligned}$$

and thus, taking absolute values of (8), we have:

$$\begin{aligned}
& |\ddot{\varphi}(t)| \\
& \leq 4 \sum_{m=1}^M \|g_m\|^2 \|F(t)f_m(t)\|^2 + 4\|\dot{F}(t)F^*(t)\|_{\text{HS}}^2 \\
& \quad + 2\|\dot{F}^*(t)F(t) + F^*(t)\dot{F}(t)\|_{\text{HS}}^2 \\
& \leq 4 \sum_{m=1}^M \|g_m\|^2 \|F(t)\|_2^2 \|f_m(t)\|^2 + 4\|\dot{F}(t)F^*(t)\|_{\text{HS}}^2 \\
& \quad + 2(\|\dot{F}^*(t)F(t)\|_{\text{HS}} + \|F^*(t)\dot{F}(t)\|_{\text{HS}})^2 \\
& \leq 4 \sum_{m=1}^M \|g_m\|^2 \|F(t)\|_{\text{HS}}^2 + 12\|\dot{F}(t)\|_{\text{HS}}^2 \|F(t)\|_{\text{HS}}^2 \\
& = 16M \sum_{m=1}^M \|g_m\|^2. \tag{9}
\end{aligned}$$

Substituting (7) and (9) into (4) yields (2). \square

In light of the Taylor expansion (2), one, in light of Cauchy's inequality, might expect the gradient of FP, namely the $\{g_m\}_{m=1}^M \in \mathbb{H}_N^M$ which maximizes the linear term

$$\text{Re} \sum_{m=1}^M \langle F^* F f_m, g_m \rangle,$$

to be given by $g_m = F^* F f_m$ for all $m = 1, \dots, M$. Indeed, one may show that this would be the correct gradient if the frame potential was being regarded as a functional over the entire space \mathbb{H}_N^M . However, as we are optimizing over \mathbb{S}_N^M , we require that $\{g_m\}_{m=1}^M \in \oplus f_m^\perp$, and as such, instead take $\{g_m\}_{m=1}^M$ to be the projection of $\{F^* F f_m\}_{m=1}^M$ onto $\oplus f_m^\perp$. In the next result, we formally verify that such a choice is indeed optimal.

Theorem 3. *For any $\{f_m\}_{m=1}^M \in \mathbb{S}_N^M$, the minimizer of the bound in (2) over all $t \in \mathbb{R}$ and $\{g_m\}_{m=1}^M \in \oplus f_m^\perp$ is given by $t = -1/(4M)$ and*

$$g_m = F^* F f_m - \|F f_m\|^2 f_m, \quad m = 1, \dots, M.$$

In particular, there exists $\{\tilde{f}_m\}_{m=1}^M \in \mathbb{S}_N^M$ such that:

$$\begin{aligned}
& \sum_{m=1}^M \|\tilde{f}_m - f_m\|^2 \\
& \leq \frac{1}{16M^2} \sum_{m=1}^M (\|F^* F f_m\|^2 - \|F f_m\|^4), \tag{10}
\end{aligned}$$

and such that:

$$\begin{aligned}
& \text{FP}(\{\tilde{f}_m\}_{m=1}^M) - \text{FP}(\{f_m\}_{m=1}^M) \\
& \leq -\frac{1}{2M} \sum_{m=1}^M (\|F^* F f_m\|^2 - \|F f_m\|^4). \tag{11}
\end{aligned}$$

Proof. We seek to minimize:

$$\begin{aligned}
& 4t \text{Re} \sum_{m=1}^M \langle F^* F f_m, g_m \rangle + 8Mt^2 \sum_{m=1}^M \|g_m\|^2 \\
& = \frac{2}{M} \sum_{m=1}^M \text{Re} \langle F^* F f_m + 2Mt g_m, 2Mt g_m \rangle \tag{12}
\end{aligned}$$

over all $\{g_m\}_{m=1}^M \in \mathbb{S}_N^M$ and all $t \in \mathbb{R}$. We note immediately from (12) that the optimal $\{g_m\}_{m=1}^M$ and t are not unique, though we now show that their product is. Indeed, for any fixed m , letting P_m denote the orthogonal projection of \mathbb{H}_N onto the orthogonal complement of f_m , we have:

$$\begin{aligned}
& \text{Re} \langle F^* F f_m + 2Mt g_m, 2Mt g_m \rangle \\
& = \text{Re} \langle F^* F f_m + 2Mt g_m, 2Mt P_m g_m \rangle \\
& = \text{Re} \langle P_m F^* F f_m + 2Mt g_m, 2Mt g_m \rangle \\
& = \frac{1}{4} (\|P_m F^* F f_m + 4Mt g_m\|^2 - \|P_m F^* F f_m\|^2) \\
& \geq -\frac{1}{4} \|P_m F^* F f_m\|^2,
\end{aligned}$$

with equality if and only if $P_m F^* F f_m + 4Mt g_m = 0$. Thus, to minimize (12), and consequently to minimize the upper bound in (2), we may take $t = -1/(4M)$ and

$$\begin{aligned}
g_m &= P_m F^* F f_m \\
&= F^* F f_m - \langle F^* F f_m, f_m \rangle f_m \\
&= F^* F f_m - \|F f_m\|^2 f_m, \tag{13}
\end{aligned}$$

as claimed. Moreover, in light of (13), we have:

$$\begin{aligned}
\|g_m\|^2 &= \langle F^* F f_m, g_m \rangle \\
&= \langle F^* F f_m, F^* F f_m - \|F f_m\|^2 f_m \rangle \\
&= \|F^* F f_m\|^2 - \|F f_m\|^4,
\end{aligned}$$

which, when substituted into (1) and (2) yields (10) and (11), respectively, where $\tilde{f}_m := f_m(-1/4M)$. \square

Note that as $\|F f_m\|^4 = |\langle F^* F f_m, f_m \rangle|^2 \leq \|F^* F f_m\|^2$ for all $m = 1, \dots, M$, Theorem 3 provides a direction and step size in which to travel from a given $\{f_m\}_{m=1}^M \in \mathbb{S}_N^M$ so as to produce a concrete decrease in frame potential. In the next section, we estimate the size of this decrease in terms of how far the current potential is from its minimum, and in so doing, provide an upper bound on the rate at which repeated applications of Theorem 3 will asymptotically produce a unit norm tight frame.

3. Gradient descent of the frame potential

We now consider the gradient descent of the frame potential: by repeatedly applying Theorem 3, we hope to produce a sequence of unit norm frames which are converging to a unit norm tight frame. Here, the main idea is to estimate the right hand side of (11) as a proportion of the difference between the current value of the frame potential and its minimum.

To be clear, in [1], the minimum value of FP over \mathbb{S}_N^M is found to be M^2/N ; we now show how the quantity $\text{FP}(\{f_m\}_{m=1}^M) - M^2/N$ is a good metric on the tightness of $\{f_m\}_{m=1}^M$. Indeed, letting $\{\lambda_n\}_{n=1}^N$ be the eigenvalues of the corresponding frame operator $F^* F$, we have:

$$\sum_{n=1}^N \lambda_n = \text{Tr}(F^* F) = \text{Tr}(F F^*) = \sum_{m=1}^M \|f_m\|^2 = M. \tag{14}$$

In particular, (14) implies that $\{f_m\}_{m=1}^M \in \mathbb{S}_N^M$ is tight if and only if $\lambda_n = \frac{M}{N}$ for all $n = 1, \dots, N$. Moreover, as

$$\text{FP}(\{f_m\}_{m=1}^M) = \|F^*F\|_{\text{HS}}^2 = \text{Tr}[(F^*F)^2] = \sum_{n=1}^N \lambda_n^2,$$

another consequence of (14) is that:

$$\begin{aligned} \text{FP}(\{f_m\}_{m=1}^M) &= \sum_{n=1}^N \left(\lambda_n - \frac{M}{N} + \frac{M}{N}\right)^2 \\ &= \sum_{n=1}^N \left(\lambda_n - \frac{M}{N}\right)^2 + 2(0) + \frac{M^2}{N}, \end{aligned}$$

and thus:

$$\text{FP}(\{f_m\}_{m=1}^M) - \frac{M^2}{N} = \sum_{n=1}^N \left(\lambda_n - \frac{M}{N}\right)^2. \quad (15)$$

That is, the difference between the frame potential and its minimum is the square of the distance of the eigenvalues of F^*F from their optimal values. Using this fact, one may show:

Theorem 4. For any $\{f_m\}_{m=1}^M \in \mathbb{S}_N^M$,

$$\begin{aligned} \sum_{m=1}^M (\|F^*F f_m\|^2 - \|F f_m\|^4) \\ \geq \delta^2 (\text{FP}(\{f_m\}_{m=1}^M) - \frac{M^2}{N}), \end{aligned} \quad (16)$$

where δ is defined as:

$$\delta := \inf_m \max_n \min_n |\langle f_m, e_n \rangle|, \quad (17)$$

where the infimum is taken over all orthonormal bases $\{e_n\}_{n=1}^N$ of \mathbb{H}_N .

For sake of space, we omit the complete proof of Theorem 4; the main idea is to let $\{e_n\}_{n=1}^N$ be an orthonormal eigenbasis of F^*F , and note that for any $m = 1, \dots, M$,

$$\begin{aligned} &\|F^*F f_m\|^2 - \|F f_m\|^4 \\ &= \left\| F^*F \sum_{n=1}^N \langle f_m, e_n \rangle e_n \right\|^2 \\ &\quad - \left| \left\langle F^*F \sum_{n=1}^N \langle f_m, e_n \rangle e_n, f_m \right\rangle \right|^2 \\ &= \left\| \sum_{n=1}^N \lambda_n \langle f_m, e_n \rangle e_n \right\|^2 - \left| \sum_{n=1}^N \langle f_m, e_n \rangle \langle \lambda_n e_n, f_m \rangle \right|^2 \\ &= \sum_{n=1}^N \lambda_n^2 |\langle f_m, e_n \rangle|^2 - \left| \sum_{n=1}^N \lambda_n |\langle f_m, e_n \rangle|^2 \right|^2 \end{aligned} \quad (18)$$

$$= \sum_{n=1}^N \left| \lambda_n - \sum_{p=1}^N \lambda_p |\langle f_m, e_p \rangle|^2 \right|^2 |\langle f_m, e_n \rangle|^2, \quad (19)$$

where the equality of (18) and (19) arises from the fact that they both represent the variance of the random variable $\{\lambda_n\}_{n=1}^N$ with respect to the probability density function $\{|\langle f_m, e_n \rangle|^2\}_{n=1}^N$.

The significance of Theorem 4 is that it bounds the decrease in frame potential given in Theorem 3 in terms of (15), that is, how far $\{f_m\}_{m=1}^M \in \mathbb{S}_N^M$ is from being tight. Indeed, using Theorem 4, one may show:

Theorem 5. For any $\{f_m\}_{m=1}^M \in \mathbb{S}_N^M$, there exists $\{\tilde{f}_m\}_{m=1}^M \in \mathbb{S}_N^M$ such that:

$$\sum_{m=1}^M \|\tilde{f}_m - f_m\|^2 \leq \frac{N+1}{16M} (\text{FP}(\{f_m\}_{m=1}^M) - \frac{M^2}{N}), \quad (20)$$

and such that:

$$\begin{aligned} \text{FP}(\{\tilde{f}_m\}_{m=1}^M) - \frac{M^2}{N} \\ \leq \left(1 - \frac{\delta^2}{2M}\right) (\text{FP}(\{f_m\}_{m=1}^M) - \frac{M^2}{N}), \end{aligned} \quad (21)$$

where δ is given in (17).

By repeatedly applying Theorem 5, one produces a sequence of unit norm frames whose tightness, measured in terms of (15), improves at a geometric rate, provided all δ 's remain above some positive lower bound; finding such a bound is a subject of current research.

4. Acknowledgments

Casazza and Fickus were supported by NSF DMS 0704216 and AFOSR F1ATA07337J001, respectively. The views expressed in this article are those of the authors and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government.

References:

- [1] J.J. Benedetto and M. Fickus. *Finite normalized tight frames*. Adv. Comput. Math., 18:357–385, 2003.
- [2] I. Bengtsson and H. Granström. *The frame potential, on average*. Preprint.
- [3] P.G. Casazza and M. Fickus. *Minimizing fusion frame potential*. To appear in Acta Appl. Math.
- [4] P.G. Casazza, M. Fickus, J. Kovačević, M. Leon and J. Tremain. A physical interpretation of tight frames. In C. Heil, editor, *Harmonic analysis and applications*, pp. 51–76, 2006.
- [5] M. Fickus, B.D. Johnson, K. Kornelson, and K. Okoudjou. *Convolutional frames and the frame potential*. Appl. Comput. Harmon. Anal., 19:77–91, 2005.
- [6] B.D. Johnson and K. Okoudjou. *Frame potential and finite abelian groups*. Contemp. Math., 464:137–148, 2008.
- [7] P. Massey. *Optimal reconstruction systems for erasures and for the q-potential*. Preprint.
- [8] P. Massey and M. Ruiz. *Minimization of convex functionals over frame operators*. To appear in Adv. Comput. Math.
- [9] P. Massey, M. Ruiz and D. Stojanoff. *The structure of minimizers of the frame potential of fusion frames*. Submitted.