

# Three proofs of the Benedetto–Fickus theorem

Dustin G. Mixon, Tom Needham, Clayton Shonkwiler, Soledad Villar

**Abstract** In 2003, Benedetto and Fickus introduced a vivid intuition for an objective function called the frame potential, whose global minimizers are fundamental objects known today as unit norm tight frames. Their main result was that the frame potential exhibits no spurious local minimizers, suggesting local optimization as an approach to construct these objects. Local optimization has since become the workhorse of cutting-edge signal processing and machine learning, and accordingly, the community has identified a variety of techniques to study optimization landscapes. This chapter applies some of these techniques to obtain three modern proofs of the Benedetto–Fickus theorem.

## 1 Introduction

Frame theory is the study of overcomplete systems of vectors in a Hilbert space. Much of the last 20 years of research in finite frame theory has concerned various modifications of the following fundamental object: (the sequence of column vectors of)  $Z = [z_1 \cdots z_n] \in \mathbb{C}^{d \times n}$  is a **unit norm tight frame** if

- $\|z_j\|^2 = 1$  for every  $j \in [n] := \{1, \dots, n\}$ , and
- $ZZ^*$  is a multiple of the  $d \times d$  identity matrix.

---

Dustin G. Mixon  
The Ohio State University, Columbus, OH, e-mail: mixon.23@osu.edu

Tom Needham  
Florida State University, Tallahassee, FL, e-mail: tneedham@fsu.edu

Clayton Shonkwiler  
Colorado State University, Fort Collins, CO, e-mail: clayton.shonkwiler@colostate.edu

Soledad Villar  
Johns Hopkins University, Baltimore, MD, e-mail: soledad.villar@jhu.edu

The multiple is necessarily the **redundancy**  $\frac{n}{d}$  of  $Z$  since  $\text{tr}(ZZ^*) = \text{tr}(Z^*Z) = n$ . Given a  $d \times n$  unit norm tight frame, it necessarily holds that

$$n \geq \text{rank}(Z) = \text{rank}(ZZ^*) = \text{rank}\left(\frac{n}{d}I_d\right) = d.$$

The unit norm tight frames with  $n = d$  are precisely the orthonormal bases for  $\mathbb{C}^d$ . When  $n > d$ , a unit norm tight frame  $Z \in \mathbb{C}^{d \times n}$  provides both an encoding  $y := Z^*x$  of data  $x \in \mathbb{C}^d$  that is robust to various types of errors [17, 10, 20] (thanks in part to the redundancy  $\frac{n}{d} > 1$ ) and a painless reconstruction formula

$$x = \frac{d}{n}ZZ^*x = \frac{d}{n}Zy$$

(thanks to the resemblance  $ZZ^* = \frac{n}{d}I_d$  to orthonormal bases).

Historically, frames were introduced in the setting of infinite-dimensional Hilbert space by Duffin and Schaeffer [13] in 1952 in the development of non-harmonic Fourier series. In 1986, Daucechies, Grossmann, and Meyer [11] further developed the theory of frames in the context of signal processing. In the time since, finite-dimensional treatments of the subject have proved interesting in their own right.

Today, certain unit norm tight frames are known to emerge as optimal frames under various objectives, for example, when hunting for  $n$  points on the unit sphere in  $\mathbb{C}^d$  for which the absolute values of the inner products between pairs of points are uniformly as small as possible [36, 30, 14]. In this chapter, we focus on an objective for which the  $d \times n$  unit norm tight frames are *precisely* the global minimizers. Let  $S(d, n)$  denote the set of matrices  $Z \in \mathbb{C}^{d \times n}$  for which every column of  $Z$  has unit norm. (We use  $S$  since each column resides in the *sphere*.) The **frame potential**  $\text{FP}: S(d, n) \rightarrow \mathbb{R}$  is defined by

$$\text{FP}(Z) := \|Z^*Z\|_F^2,$$

where  $\|\cdot\|_F$  denotes the Frobenius matrix norm.

**Proposition 1 (Welch [36])** *For every  $Z \in S(d, n)$ , it holds that  $\text{FP}(Z) \geq \frac{n^2}{d}$ , with equality precisely when  $Z$  is a unit norm tight frame.*

**Proof** Rearrange the inequality

$$0 \leq \|ZZ^* - \frac{n}{d}I_d\|_F^2 = \|ZZ^*\|_F^2 - 2\langle ZZ^*, \frac{n}{d}I_d \rangle + \|\frac{n}{d}I_d\|_F^2 = \text{FP}(Z) - \frac{n^2}{d},$$

and observe that equality holds precisely when  $ZZ^* = \frac{n}{d}I_d$ .  $\square$

We note that when  $n < d$ , Welch's bound is dominated by the simpler bound

$$\text{FP}(Z) \geq n,$$

and equality is achieved precisely when the columns of  $Z$  are orthonormal.

In [4], Benedetto and Fickus introduced the idea of locally minimizing the frame potential, much like how point charges confined to a sphere will dynamically repel

each other in order to locally minimize a Coulomb potential. For context, there are several longstanding open problems that concern minimizing a given potential of a point configuration on the sphere, to include the Thomson problem [35], the Tammes problem [33], and Smale’s 7th problem [29]. The difficulty with these problems stems from what appear to be exponentially many spurious local minimizers, i.e., local minimizers that are not global minimizers. Perhaps surprisingly, the frame potential does not share this defect:

**Proposition 2 (Benedetto–Fickus theorem [4])** *For each  $d, n \in \mathbb{N}$ , the frame potential  $FP: S(d, n) \rightarrow \mathbb{R}$  has no spurious local minimizers.*

We offer two comments about the results in [4]. First, Benedetto and Fickus consider both the real and complex cases. We isolate the complex case here because it appears to be more amenable to analysis by a variety of interesting techniques. Throughout, we discuss to what extent certain arguments can be transferred to the real case. Second, even if we were to include the real case, Benedetto and Fickus prove a bit more than what is contained in the Benedetto–Fickus theorem. Specifically, they use Lagrange multipliers to characterize the critical points of the frame potential, and then they find a descent direction for every critical point that is neither orthonormal (in the case where  $n \leq d$ ) nor a unit norm tight frame (in the case where  $n \geq d$ ). Note that by compactness, a global minimum necessarily exists, which in turn is a critical point with no descent direction, and by the above, this point must be a unit norm tight frame whenever  $n \geq d$ . By this argument, Benedetto and Fickus provide a clever implicit proof of the existence of unit norm tight frames and the achievability of Welch’s bound.

Regardless, we isolate the Benedetto–Fickus theorem as the main result of [4]. One reason is that this feature of the frame potential clearly distinguishes it from other potentials such as the Coulomb potential. Second, there are much shorter explicit proofs of the existence of unit norm tight frames (even in the real case): for each  $d, n \in \mathbb{N}$  with  $n \geq d$ , it suffices to select  $d$  appropriate rows of the  $n \times n$  real discrete Fourier transform matrix and scale the resulting columns to have unit norm. Indeed, the frame potential is unusual in how easy it is to characterize its global minimizers. Finally, our statement of the Benedetto–Fickus theorem highlights the relevance of [4] to modern trends in signal processing and machine learning. In particular, much of the last decade of research has centered around analyzing non-convex landscapes for optimization in various settings, such as orthogonal tensor decomposition [15], the Burer–Monteiro method [6], dictionary learning [31], matrix completion [16], phase retrieval [32], and neural network training [23]. In all of these settings, a primary goal has been to establish the absence of spurious local minimizers so as to explain the empirical performance of local optimization.

The recent literature has introduced many techniques to our community, and in this chapter, we use some of them to prove three different strengthenings of the Benedetto–Fickus theorem. The proof in the next section is closest to the original, though we use a stronger Lagrange multiplier result to make the proof more routine, and we use Wirtinger derivatives to simplify the necessary calculations. In Section 3, we apply a general observation that a convex function of an open map has no spurious

local minimizers, and we obtain such a factorization of the frame potential using the theory of *eigensteps* developed in [7]. Finally, Section 4 leverages ideas from geometric invariant theory to show that locally minimizing the frame potential by gradient flow sends almost every initialization to a unit norm tight frame.

## 2 Lagrange multipliers and Wirtinger derivatives

Lagrange multipliers provide a standard approach for optimization subject to equality constraints. What follows is a fundamental result in this setting:

**Proposition 3 (Lagrange multiplier theorem, Proposition 3.1.1 in [5])** *Given twice continuously differentiable functions  $f, h_1, \dots, h_m: \mathbb{R}^n \rightarrow \mathbb{R}$ , consider*

$$S := \{x \in \mathbb{R}^n : h_1(x) = \dots = h_m(x) = 0\},$$

*and suppose  $x_0 \in S$  is a local minimum of the restriction of  $f$  to  $S$  with linearly independent constraint gradients  $\{\nabla h_i(x_0)\}_{i \in [m]}$ . Then there exists a unique choice of Lagrange multipliers  $\lambda_1, \dots, \lambda_m \in \mathbb{R}$  such that*

$$\nabla f(x_0) + \sum_{i \in [m]} \lambda_i \nabla h_i(x_0) = 0, \quad y^\top \left( \nabla^2 f(x_0) + \sum_{i \in [m]} \lambda_i \nabla^2 h_i(x_0) \right) y \geq 0 \quad (1)$$

*for all  $y \in (\text{span}\{\nabla h_i(x_0)\}_{i \in [m]})^\perp$ .*

This motivates a couple of definitions. In the setting of Proposition 3, we say  $x_0 \in S$  is a **(first-order) critical point** if the constraint gradients  $\{\nabla h_i(x_0)\}_{i \in [m]}$  are linearly dependent or there exist Lagrange multipliers  $\lambda_1, \dots, \lambda_m \in \mathbb{R}$  such that

$$\nabla f(x_0) + \sum_{i \in [m]} \lambda_i \nabla h_i(x_0) = 0,$$

and we say  $x_0 \in S$  is a **second-order critical point** if  $\{\nabla h_i(x_0)\}_{i \in [m]}$  are linearly dependent or there exist  $\lambda_1, \dots, \lambda_m \in \mathbb{R}$  such that (1) holds for all  $y \in (\text{span}\{\nabla h_i(x_0)\}_{i \in [m]})^\perp$ . By Proposition 3, every local minimizer is a second-order critical point, and by relaxation, every second-order critical point is a first-order critical point. (Indeed, we include the dependence of constraint gradients in the definition of critical points so that this implication holds, but as we will see, such degeneracy never arises in our setting.)

In this section, we simplify the original proof of the Benedetto–Fickus theorem to obtain a proof of the following strengthening:

**Theorem 1** *For each  $d, n \in \mathbb{N}$ , the second-order critical points of the frame potential  $\text{FP}: S(d, n) \rightarrow \mathbb{R}$  are precisely the global minimizers of the frame potential, namely, the  $d \times n$  matrices with orthonormal columns when  $n \leq d$ , and the  $d \times n$  unit norm tight frames when  $n \geq d$ .*

While we will prove Theorem 1 in the complex setting, the same proof works (and is even simpler) in the real setting. In our setting, we are optimizing the restriction of a function  $f: \mathbb{C}^n \rightarrow \mathbb{R}$ , and so to find critical points, we are inclined to take  $g: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{C}^n$  defined by  $g(x, y) := x + iy$  and then compute the gradient of the composition  $f \circ g$ . Working with this composition can be cumbersome, but as we will see, Wirtinger derivatives nicely encode the desired gradient information in terms of functions over the native space  $\mathbb{C}^n$ . This formalism drew widespread attention from the signal processing community when Candès, Li, and Soltanolkotabi used it to analyze the phase retrieval problem in [9]. The **Wirtinger derivative**  $\frac{\partial f}{\partial \bar{z}_i}: \mathbb{C}^n \rightarrow \mathbb{C}$  is defined by

$$\frac{\partial f}{\partial \bar{z}_i} := \frac{1}{2} \left( \frac{\partial(f \circ g)}{\partial x_i} + i \frac{\partial(f \circ g)}{\partial y_i} \right) \circ g^{-1},$$

which in turn determines the **Wirtinger gradient**  $\frac{\partial f}{\partial \bar{z}}: \mathbb{C}^n \rightarrow \mathbb{C}^n$  defined by

$$\frac{\partial f}{\partial \bar{z}} := \left( \frac{\partial f}{\partial \bar{z}_1}, \dots, \frac{\partial f}{\partial \bar{z}_n} \right).$$

For example, given  $f: \mathbb{C}^n \rightarrow \mathbb{R}$  defined by  $f(z) := \|z\|^2$ , we have  $(f \circ g)(x, y) = \sum_{j \in [n]} (x_j^2 + y_j^2)$ , and so  $\frac{\partial(f \circ g)}{\partial x_i}(x, y) = 2x_i$ ,  $\frac{\partial(f \circ g)}{\partial y_i}(x, y) = 2y_i$ , and  $\frac{\partial f}{\partial \bar{z}}(z) = z$ . Next, the **gradient**  $\nabla f: \mathbb{C}^n \rightarrow \mathbb{C}^n$  is defined in terms of the best real linear approximation of  $f$  at the input  $z$ :

$$f(z + \Delta z) = f(z) + \operatorname{Re}\langle \nabla f(z), \Delta z \rangle + \text{h.o.t.}$$

Since  $g$  is a real linear isometry, it follows that  $\nabla(f \circ g) = g^{-1} \circ \nabla f \circ g$ , meaning  $\nabla f$  is the notion of gradient that is relevant to optimization. One may verify that

$$\nabla f = 2 \frac{\partial f}{\partial \bar{z}}.$$

For example, for  $f(z) := \|z\|^2$ , we have  $\nabla f(z) = 2z$ , which nicely resembles the gradient expression when  $z$  is real. The most attractive aspect of Wirtinger derivatives is that if  $f(z)$  is a polynomial in  $z$  and  $\bar{z}$ , then  $\frac{\partial f}{\partial \bar{z}_i}(z)$  may be computed by formally differentiating in  $\bar{z}_i$  as if everything else (including  $z_i$ ) is constant. For example, for  $f(z) := \|z\|^2 = \sum_{j \in [n]} z_j \bar{z}_j$ , we have  $\frac{\partial f}{\partial \bar{z}_i}(z) = z_i$ , matching our previous calculation.

**Lemma 1** Consider the extension  $\text{EP}: \mathbb{C}^{d \times n} \rightarrow \mathbb{R}$  of the frame potential defined by  $\text{EP}(Z) := \|Z^* Z\|_F^2$ , and for each  $j \in [n]$ , consider the constraint  $h_j: \mathbb{C}^{d \times n} \rightarrow \mathbb{R}$  defined by  $h_j(Z) := \|Z e_j\|^2 - 1$ , where  $e_j$  denotes the  $j$ th standard basis element in  $\mathbb{C}^n$ . Then for  $Z = [z_1 \cdots z_n]$ , it holds that

$$\nabla \text{EP}(Z) = 4ZZ^*Z, \quad \nabla h_j(Z) = 2z_j e_j^*.$$

**Proof** For EP, we expand

$$\text{EP}(Z) = \|Z^*Z\|_F^2 = \sum_{i \in [n]} \sum_{j \in [n]} |(Z^*Z)_{ij}|^2 = \sum_{i \in [n]} \sum_{j \in [n]} \sum_{k \in [d]} \sum_{\ell \in [d]} \bar{Z}_{ki} Z_{kj} Z_{\ell i} \bar{Z}_{\ell j},$$

and formally differentiate

$$\begin{aligned} \frac{\partial \text{EP}}{\partial \bar{Z}_{pq}}(Z) &= \sum_{i \in [n]} \sum_{j \in [n]} \sum_{k \in [d]} \sum_{\ell \in [d]} (\delta_{pk} \delta_{qi} Z_{kj} Z_{\ell i} \bar{Z}_{\ell j} + \bar{Z}_{ki} Z_{kj} Z_{\ell i} \delta_{p\ell} \delta_{qj}) \\ &= \sum_{j \in [n]} \sum_{\ell \in [d]} Z_{pj} Z_{\ell q} \bar{Z}_{\ell j} + \sum_{i \in [n]} \sum_{k \in [d]} \bar{Z}_{ki} Z_{kq} Z_{pi} \\ &= \sum_{j \in [n]} \sum_{\ell \in [d]} Z_{pj} (Z^*)_{j\ell} Z_{\ell q} + \sum_{i \in [n]} \sum_{k \in [d]} Z_{pi} (Z^*)_{ik} Z_{kq} = (ZZZ^*Z)_{pq}, \end{aligned}$$

and then double the Wirtinger gradient to obtain the gradient. For  $h_j$ , we similarly expand and formally differentiate

$$h_j(Z) = \sum_{i \in [d]} Z_{ij} \bar{Z}_{ij} - 1, \quad \frac{\partial h_j}{\partial \bar{Z}_{pq}}(Z) = \sum_{i \in [d]} Z_{ij} \delta_{pi} \delta_{qj} = Z_{pj} \delta_{qj} = (z_j e_j^*)_{pq},$$

and then double the Wirtinger gradient to obtain the gradient.  $\square$

**Proposition 4** *For each  $d, n \in \mathbb{N}$ , the first-order critical points of the frame potential  $\text{FP}: \mathcal{S}(d, n) \rightarrow \mathbb{R}$  are precisely the matrices  $Z \in \mathcal{S}(d, n)$  for which every column of  $Z$  is an eigenvector of  $ZZ^*$ .*

**Proof** For every  $Z \in \mathcal{S}(d, n)$ , it holds that  $\{\nabla h_j(Z)\}_{j \in [n]}$  are nonzero matrices with disjoint support, meaning they are linearly independent. Thus,  $Z$  is a first-order critical point if and only if  $\nabla \text{EP}(Z)$  is in the span of  $\{\nabla h_j(Z)\}_{j \in [n]}$ , if and only if for every  $j \in [n]$ , the  $j$ th column of  $\nabla \text{EP}(Z)$  is a scalar multiple of the  $j$ th column of  $\nabla h_j(Z)$ . The result follows.  $\square$

Now that we have characterized the first-order critical points of the frame potential, we isolate a useful feature of the non-minimizing critical points. To enunciate this feature, we need another definition: We say  $Z \in \mathcal{S}(d, n)$  is a **unit norm tight frame for a subspace**  $S \subseteq \mathbb{C}^d$  if  $ZZ^*$  is a multiple of the orthogonal projection matrix with image  $S$ . In particular, the multiple equals  $n/\dim(S)$  by a trace argument.

**Corollary 1** *Fix  $d, n \in \mathbb{N}$ , and suppose  $Z \in \mathcal{S}(d, n)$  is a first-order critical point of the frame potential  $\text{FP}: \mathcal{S}(d, n) \rightarrow \mathbb{R}$  that is not a global minimizer. Then there exist  $d', n' \in \mathbb{N}$  satisfying*

$$d' < d, \quad \frac{n'}{d'} > \max\left\{\frac{n}{d}, 1\right\}$$

*such that the maximum eigenvalue of  $ZZ^*$  has a  $d'$ -dimensional eigenspace  $E$  that contains exactly  $n'$  columns of  $Z$ , which in turn form a unit norm tight frame for  $E$ .*

**Proof** Let  $\Lambda$  denote the set of eigenvalues of  $ZZ^*$ , and let  $d'$  denote the multiplicity of  $\max \Lambda$  as an eigenvalue of  $ZZ^*$ . Since  $Z$  is not a unit norm tight frame by assumption,

it holds that  $d' < d$ . Next, let  $P_\lambda$  denote the orthogonal projection matrix onto the eigenspace of  $ZZ^*$  corresponding to  $\lambda \in \Lambda$ . Then the spectral theorem gives

$$\sum_{\lambda \in \Lambda} \lambda \frac{\text{tr} P_\lambda}{d} = \frac{1}{d} \text{tr} \left( \sum_{\lambda \in \Lambda} \lambda P_\lambda \right) = \frac{1}{d} \text{tr}(ZZ^*) = \frac{1}{d} \text{tr}(Z^*Z) = \frac{n}{d},$$

i.e.,  $\frac{n}{d}$  is a weighted average of the eigenvalues of  $ZZ^*$ . Since  $Z$  is not a unit norm tight frame by assumption, there are at least two distinct eigenvalues, and so  $\max \Lambda > \frac{n}{d}$ . In the case where  $n < d$ , the analogous weighted average of eigenvalues of  $Z^*Z$  equals 1, and so  $\max \Lambda > 1$  follows from the assumption that the columns of  $Z$  are not orthonormal. Overall, we have

$$d' < d, \quad \max \Lambda > \max \left\{ \frac{n}{d}, 1 \right\}.$$

Next, consider the assignment  $f: [n] \rightarrow \Lambda$  afforded by Proposition 4. Then the spectral theorem gives

$$\sum_{\lambda \in \Lambda} \sum_{j \in f^{-1}(\lambda)} z_j z_j^* = \sum_{j \in [n]} z_j z_j^* = ZZ^* = \sum_{\lambda \in \Lambda} \lambda P_\lambda.$$

Since distinct eigenspaces are orthogonal, then for each  $\lambda' \in \Lambda$ , we have

$$\sum_{j \in f^{-1}(\lambda')} z_j z_j^* = P_{\lambda'} \sum_{\lambda \in \Lambda} \sum_{j \in f^{-1}(\lambda)} z_j z_j^* = P_{\lambda'} \sum_{\lambda \in \Lambda} \lambda P_\lambda = \lambda' P_{\lambda'},$$

i.e.,  $\{z_j z_j^*\}_{j \in f^{-1}(\lambda')}$  is a unit norm tight frame for the eigenspace corresponding to  $\lambda'$ . Taking  $n' := |f^{-1}(\max \Lambda)|$  gives  $\max \Lambda = \frac{n'}{d}$ , and the result follows.  $\square$

Next, we will characterize the second-order critical points of the frame potential. To accomplish this, we first recall how Wirtinger calculus captures Hessian information. The Wirtinger derivative  $\frac{\partial f}{\partial z_i}: \mathbb{C}^n \rightarrow \mathbb{C}$  is defined by

$$\frac{\partial f}{\partial z_i} := \frac{1}{2} \left( \frac{\partial(f \circ g)}{\partial x_i} - i \frac{\partial(f \circ g)}{\partial y_i} \right) \circ g^{-1},$$

and the **Wirtinger Hessians**  $\frac{\partial^2 f}{\partial z^2}, \frac{\partial^2 f}{\partial z \partial \bar{z}}, \frac{\partial^2 f}{\partial \bar{z} \partial z}, \frac{\partial^2 f}{\partial \bar{z}^2}: \mathbb{C}^n \rightarrow \mathbb{C}^{n \times n}$  are defined by

$$\left( \frac{\partial^2 f}{\partial z \partial \bar{z}}(u) \right)_{ij} := \frac{\partial}{\partial z_i} \left( \frac{\partial f}{\partial \bar{z}_j} \right)(u),$$

and similarly for the other Wirtinger Hessians. The following result is essentially contained in the monograph [22], but the proof is straightforward, so we provide a sketch for completeness:

**Proposition 5** Define  $g: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{C}^n$  by  $g(x, y) := x + iy$  and  $h: \mathbb{C}^n \rightarrow \mathbb{C}^n \times \mathbb{C}^n$  by  $h(z) := (z, \bar{z})$ . Consider any  $f: \mathbb{C}^n \rightarrow \mathbb{R}$  such that  $f \circ g$  is twice continuously

differentiable. For every  $u, v \in \mathbb{R}^n \times \mathbb{R}^n$ , it holds that

$$v^\top \nabla^2(f \circ g)(u)v = (h \circ g)(v)^\top \begin{bmatrix} \frac{\partial^2 f}{\partial z^2}(g(u)) & \frac{\partial^2 f}{\partial z \partial \bar{z}}(g(u)) \\ \frac{\partial^2 f}{\partial \bar{z} \partial z}(g(u)) & \frac{\partial^2 f}{\partial \bar{z}^2}(g(u)) \end{bmatrix} (h \circ g)(v).$$

**Proof (sketch)** By a standard abuse of notation, we write

$$\frac{\partial}{\partial x_i} = \frac{\partial}{\partial z_i} + \frac{\partial}{\partial \bar{z}_i}, \quad \frac{\partial}{\partial y_i} = i \left( \frac{\partial}{\partial z_i} - \frac{\partial}{\partial \bar{z}_i} \right).$$

Then, for example,

$$\frac{\partial^2}{\partial y_i \partial x_j} = i \left( \frac{\partial}{\partial z_i} - \frac{\partial}{\partial \bar{z}_i} \right) \left( \frac{\partial}{\partial z_j} + \frac{\partial}{\partial \bar{z}_j} \right),$$

and expanding implies

$$\nabla_{yx}^2 = i \left( \frac{\partial^2}{\partial z^2} + \frac{\partial^2}{\partial z \partial \bar{z}} - \frac{\partial^2}{\partial \bar{z} \partial z} - \frac{\partial^2}{\partial \bar{z}^2} \right).$$

The same approach delivers similar expressions for  $\nabla_{xx}^2$ ,  $\nabla_{xy}^2$ , and  $\nabla_{yy}^2$ . The desired result can then be verified by expanding the left-hand side in terms of these expressions and rearranging to obtain the right-hand side.  $\square$

Proposition 5 will help us identify the second-order critical points of the frame potential. To this end, the following lemma summarizes a straightforward calculation:

**Lemma 2** For EP,  $h_1, \dots, h_n : \mathbb{C}^{d \times n} \rightarrow \mathbb{R}$  defined in Lemma 1, it holds that

$$\frac{\partial^2 \text{EP}}{\partial Z_{pq} \partial Z_{rs}}(Z) = 2\bar{Z}_{rq} \bar{Z}_{ps}, \quad \frac{\partial^2 \text{EP}}{\partial \bar{Z}_{pq} \partial \bar{Z}_{rs}}(Z) = 2Z_{rq} Z_{ps},$$

$$\frac{\partial^2 \text{EP}}{\partial Z_{pq} \partial \bar{Z}_{rs}}(Z) = 2 \left( \delta_{pr} (Z^* Z)_{qs} + \delta_{qs} (Z Z^*)_{rp} \right),$$

$$\frac{\partial^2 h_j}{\partial Z_{pq} \partial Z_{rs}}(Z) = \frac{\partial^2 h_j}{\partial \bar{Z}_{pq} \partial \bar{Z}_{rs}}(Z) = 0, \quad \frac{\partial^2 h_j}{\partial Z_{pq} \partial \bar{Z}_{rs}}(Z) = \delta_{pr} \delta_{qj} \delta_{sj}.$$

We may now prove the main result of this section:

**Proof (of Theorem 1)** One containment is trivial: every global minimizer is necessarily a local minimizer, which is necessarily a second-order critical point. For the other direction, we show that every point  $Z \in \mathcal{S}(d, n)$  that is not a global minimizer is necessarily not a second-order critical point. Since every second-order critical point is necessarily a first-order critical point, the implication trivially holds if  $Z$  is not a first-order critical point. Suppose  $Z = [z_1 \cdots z_n]$  is a first-order critical point that is not a global minimizer. The corresponding Lagrange multipliers  $\{\lambda_j\}_{j \in [n]}$  are the unique solution to



$$ZZ^* z_j = -\frac{1}{2} \lambda_j z_j$$

for  $j \in [n]$ . Consider the map  $L: \mathbb{C}^{d \times n} \rightarrow \mathbb{R}$  defined by

$$L(Z) := \text{EP}(Z) + \sum_{j \in [n]} \lambda_j h_j(Z).$$

To show that  $Z$  is not a second-order critical point, Proposition 5 gives that it suffices to find  $Y \in \mathbb{C}^{d \times n}$  such that

$$\text{Re tr}(\nabla h_j(Z)^* Y) = 0$$

for every  $j \in [n]$  and

$$\begin{aligned} \sum_{p \in [d]} \sum_{q \in [n]} \sum_{r \in [d]} \sum_{s \in [n]} \left( \frac{\partial^2 L(Z)}{\partial Z_{pq} \partial Z_{rs}} Y_{pq} Y_{rs} + \frac{\partial^2 L(Z)}{\partial Z_{pq} \partial \bar{Z}_{rs}} Y_{pq} \bar{Y}_{rs} \right. \\ \left. + \frac{\partial^2 L(Z)}{\partial \bar{Z}_{pq} \partial Z_{rs}} \bar{Y}_{pq} Y_{rs} + \frac{\partial^2 L(Z)}{\partial \bar{Z}_{pq} \partial \bar{Z}_{rs}} \bar{Y}_{pq} \bar{Y}_{rs} \right) < 0. \end{aligned}$$

To enunciate our choice for  $Y$ , let  $J \subseteq [n]$  denote the index set of the  $n'$  columns of  $Z$  implicated by Corollary 1. Then Corollary 1 gives that  $\{z_j\}_{j \in J}$  spans a  $d'$ -dimensional eigenspace  $E$  of  $ZZ^*$  with  $d' < \min\{d, n'\}$ . Since  $d' < d$ , there is another eigenspace  $E'$  of  $ZZ^*$ , and we select  $u \in E'$  with  $\|u\|^2 = 1$ . Since  $d' < n'$ , then  $\{z_j\}_{j \in J}$  is linearly dependent, and we select  $v \in \mathbb{C}^n$  with support contained in  $J$  such that  $Zv = 0$  and  $\|v\|^2 = 1$ . We take  $Y := uv^*$ , and it remains to verify that  $Y$  satisfies the above conditions. First, Lemma 1 gives

$$\text{Re tr}(\nabla h_j(Z)^* Y) = \text{Re tr}(2e_j z_j^* u v^*) = 2 \text{Re}(v^* e_j z_j^* u) = 0,$$

where the last step follows from the facts that  $v^* e_j = 0$  whenever  $j \notin J$  and  $z_j^* u = 0$  whenever  $j \in J$ . Next, Lemma 2 gives

$$\begin{aligned} \sum_{p \in [d]} \sum_{q \in [n]} \sum_{r \in [d]} \sum_{s \in [n]} \frac{\partial^2 \text{EP}(Z)}{\partial Z_{pq} \partial Z_{rs}} Y_{pq} Y_{rs} &= \sum_{p \in [d]} \sum_{q \in [n]} \sum_{r \in [d]} \sum_{s \in [n]} 2 \bar{Z}_{rq} \bar{Z}_{ps} u_p \bar{v}_q u_r \bar{v}_s \\ &= 2(v^* Z^* u)^2 = 0, \end{aligned}$$

where the last step follows from the fact that  $v$  and  $Z^* u$  have disjoint support. We similarly have

$$\sum_{p \in [d]} \sum_{q \in [n]} \sum_{r \in [d]} \sum_{s \in [n]} \frac{\partial^2 \text{EP}(Z)}{\partial \bar{Z}_{pq} \partial \bar{Z}_{rs}} \bar{Y}_{pq} \bar{Y}_{rs} = 0.$$

Considering Lemma 2, it remains to show that

$$\text{Re} \left( \sum_{p \in [d]} \sum_{q \in [n]} \sum_{r \in [d]} \sum_{s \in [n]} \frac{\partial^2 L(Z)}{\partial Z_{pq} \partial \bar{Z}_{rs}} Y_{pq} \bar{Y}_{rs} \right) < 0. \quad (2)$$

To this end, let  $\alpha$  and  $\beta$  denote the eigenvalues corresponding to eigenspaces  $E$  and  $E'$ , respectively. Notably, Corollary 1 gives  $\alpha > \beta$ , and this will be the source of our desired inequality. Lemma 2 gives

$$\begin{aligned}
& \sum_{p \in [d]} \sum_{q \in [n]} \sum_{r \in [d]} \sum_{s \in [n]} \frac{\partial^2 \text{EP}(Z)}{\partial Z_{pq} \partial \bar{Z}_{rs}} Y_{pq} \bar{Y}_{rs} \\
&= \sum_{p \in [d]} \sum_{q \in [n]} \sum_{r \in [d]} \sum_{s \in [n]} 2 \left( \delta_{pr} (Z^* Z)_{qs} + \delta_{qs} (ZZ^*)_{rp} \right) u_p \bar{v}_q \bar{u}_r v_s \\
&= 2 \sum_{q \in [n]} \sum_{r \in [d]} \sum_{s \in [n]} (Z^* Z)_{qs} u_r \bar{v}_q \bar{u}_r v_s + 2 \sum_{p \in [d]} \sum_{r \in [d]} \sum_{s \in [n]} (ZZ^*)_{rp} u_p \bar{v}_s \bar{u}_r v_s \\
&= 2 \|Zv\|^2 \|u\|^2 + 2 \|Z^* u\|^2 \|v\|^2 = 2\beta, \tag{3}
\end{aligned}$$

where the last step uses the facts that  $Zv = 0$ ,  $\|v\|^2 = 1$ , and  $\|Z^* u\|^2 = u^* Z Z^* u = \beta$ . Next, Lemma 2 gives

$$\begin{aligned}
\sum_{p \in [d]} \sum_{q \in [n]} \sum_{r \in [d]} \sum_{s \in [n]} \frac{\partial^2 h_j(Z)}{\partial Z_{pq} \partial \bar{Z}_{rs}} Y_{pq} \bar{Y}_{rs} &= \sum_{p \in [d]} \sum_{q \in [n]} \sum_{r \in [d]} \sum_{s \in [n]} \delta_{pr} \delta_{qj} \delta_{sj} u_p \bar{v}_q \bar{u}_r v_s \\
&= \|u\|^2 |v_j|^2 = |v_j|^2. \tag{4}
\end{aligned}$$

We combine (3) and (4) to get

$$\sum_{p \in [d]} \sum_{q \in [n]} \sum_{r \in [d]} \sum_{s \in [n]} \frac{\partial^2 L(Z)}{\partial Z_{pq} \partial \bar{Z}_{rs}} Y_{pq} \bar{Y}_{rs} = 2\beta + \sum_{j \in [n]} \lambda_j |v_j|^2 = 2(\beta - \alpha) < 0,$$

where the last equality follows from the facts that  $\lambda_j = -2\alpha$  whenever  $j \in J$ ,  $v_j = 0$  whenever  $j \notin J$ , and  $\|v\|^2 = 1$ . This implies the desired inequality (2).  $\square$

### 3 Convexity and topology

Given a self-adjoint matrix  $A \in \mathbb{C}^{d \times d}$ , let  $\lambda(A) \in \mathbb{R}^d$  denote the vector of eigenvalues of  $A$  sorted in monotonically decreasing order. This section proves the following strengthening of the Benedetto–Fickus theorem:

**Theorem 2** *Given a convex function  $f: \mathbb{R}^d \rightarrow \mathbb{R}$ , consider  $\text{FP}_f: \mathcal{S}(d, n) \rightarrow \mathbb{R}$  defined by  $\text{FP}_f(Z) := f(\lambda(ZZ^*))$ . Then  $\text{FP}_f$  has no spurious local minimizers.*

Notably, if  $f(\cdot) = \|\cdot\|^2$ , then  $\text{FP}_f$  is the frame potential. We will prove Theorem 2 using ideas from both convexity and topology. Recall that given topological spaces  $(X, \tau_X)$  and  $(Y, \tau_Y)$ , we say  $f: X \rightarrow Y$  is **open at**  $x \in X$  if for every  $U \in \tau_X$  such that  $x \in U$ , there exists  $V \in \tau_Y$  such that  $f(x) \in V \subseteq f(U)$ . Furthermore,  $f$  is **open** if for every  $x \in X$ ,  $f$  is open at  $x$ . Throughout this section,  $\text{int}(U)$  and  $\text{cl}(U)$  denote

the topological interior and closure of  $U$ , respectively. The following observation appears to originate with the recent preprint [28]:

**Lemma 3** *Given a topological space  $X$ , a convex subset  $Y$  of a real topological vector space, an open map  $g: X \rightarrow Y$  in the subspace topology for  $Y$ , and a convex function  $f: Y \rightarrow \mathbb{R}$ , it holds that  $f \circ g$  has no spurious local minimizers.*

**Proof** Suppose  $x_0 \in X$  is a local minimizer of  $f \circ g$ , i.e., there exists  $U \in \tau_X$  with  $x_0 \in U$  such that  $(f \circ g)(x) \geq (f \circ g)(x_0)$  for every  $x \in U$ . Since  $g$  is open, it follows that  $V := g(U) \in \tau_Y$  contains  $g(x_0)$  and has the property that  $f(y) \geq f(g(x_0))$  for every  $y \in V$ . That is,  $g(x_0)$  is a local minimizer of  $f$ . Since  $f$  is convex,  $g(x_0)$  is necessarily a global minimizer of  $f$ , meaning  $f(y) \geq f(g(x_0))$  for every  $y \in Y$ . Then  $f(g(x)) \geq f(g(x_0))$  for every  $x \in X$ , i.e.,  $x_0$  is a global minimizer of  $f \circ g$ .  $\square$

As such, to prove Theorem 2, it suffices to factor  $\text{FP}_f$  as the composition of an open map with a convex function. To proceed, we need a definition. Given  $x, y \in \mathbb{R}^d$ , we say  $x$  **interlaces**  $y$  and write  $x \sqsupseteq y$  if

$$x_1 \geq y_1 \geq x_2 \geq y_2 \geq \cdots \geq x_d \geq y_d.$$

Let  $E(d, n)$  denote the polytope of matrices  $[\lambda_1 \cdots \lambda_n] \in \mathbb{R}^{d \times n}$  such that

$$\lambda_n \sqsupseteq \cdots \sqsupseteq \lambda_1 \sqsupseteq 0 \quad \text{and} \quad \sum_{i \in [d]} (\lambda_j)_i = j \quad \forall j \in [n].$$

For example, one may verify that

$$\frac{1}{4} \begin{bmatrix} 4 & 6 & 7 & 8 & 9 \\ 0 & 2 & 4 & 6 & 7 \\ 0 & 0 & 1 & 2 & 4 \end{bmatrix} \in E(3, 5).$$

Throughout this section, we use  $\Lambda$  to denote a member of  $E(d, n)$ .

Define  $\Sigma: S(d, n) \rightarrow E(d, n)$  so that the  $j$ th column of  $\Sigma([z_1 \cdots z_n])$  is the vector of eigenvalues of  $\sum_{i \in [j]} z_i z_i^*$  sorted in monotonically decreasing order. The fact that  $\Sigma(Z) \in E(d, n)$  for every  $Z \in S(d, n)$  essentially follows from Cauchy’s interlacing theorem and is established in the proof of the  $(\Rightarrow)$  direction of Theorem 2 in [7]. In words,  $\Sigma(Z)$  is known as the **eigensteps** of  $Z$ . (We use  $\Sigma$  since it looks like an E for *eigen* and sounds like an S for *steps*.) Observe the relationship between eigensteps and Gelfand–Tsetlin polytopes [12]. Eigensteps were introduced in [7] to explicitly construct the set of unit norm tight frames, and were later used to establish the path-connectivity of this set [8].

Considering  $\lambda(ZZ^*) = \Sigma(Z)e_d$ , then  $\text{FP}_f$  is the composition of  $Z \mapsto \Sigma(Z)$  with the convex function  $\Lambda \mapsto f(\Lambda e_d)$ . Thus, we would like to show that the eigensteps map  $\Sigma$  is open. The following lemma summarizes our approach to this end:

**Lemma 4** *Given topological spaces  $X$  and  $Y$ , consider a point  $x \in X$  and a map  $f: X \rightarrow Y$  that is continuous at  $x$ . Suppose there exists  $W \in \tau_Y$  such that  $f(x) \in W$  and a continuous map  $g: W \rightarrow X$  such that*

$$(g \circ f)(x) = x \quad \text{and} \quad f \circ g = \text{id}_W .$$

Then  $f$  is open at  $x$ .

**Proof** Given any  $U \in \tau_X$  such that  $x \in U$ , take  $V := g^{-1}(U \cap f^{-1}(W))$ . Then  $V \in \tau_Y$  since  $W$  is open,  $f(x) \in W$ ,  $f$  is continuous at  $x$ ,  $U$  is open, and  $g$  is continuous. Next, since  $x \in U$  and  $f(x) \in W$ , we have  $x \in U \cap f^{-1}(W)$ , and so  $g(f(x)) = (g \circ f)(x) = x$  gives  $f(x) \in g^{-1}(U \cap f^{-1}(W)) = V$ . It remains to verify that  $V \subseteq f(U)$ . Select any  $v \in V = g^{-1}(U \cap f^{-1}(W)) \subseteq W$ . Then  $v = (f \circ g)(v) = f(g(v)) \in f(U \cap f^{-1}(W)) \subseteq f(U)$ .  $\square$

Due to the nature of the eigensteps map, our construction of  $g$  uses a bit of matrix analysis. We start with a couple of lemmas:

**Lemma 5** *Given a self-adjoint matrix  $A \in \mathbb{C}^{d \times d}$  and eigenvalue  $\lambda$  of multiplicity  $m$ , let  $P_\lambda \in \mathbb{C}^{d \times d}$  denote orthogonal projection onto the corresponding eigenspace. For  $z \in \mathbb{C}^d$  with  $P_\lambda z \neq 0$ , the multiplicity of  $\lambda$  as an eigenvalue of  $A + zz^*$  is  $m - 1$ .*

**Proof** Let  $\{v_i\}_{i \in [n]}$  denote any orthonormal basis of eigenvectors of  $A$  with corresponding eigenvalues  $\{\lambda_i\}_{i \in [n]}$ . We apply the matrix determinant lemma to express the characteristic polynomial of  $A + zz^*$ :

$$\begin{aligned} p_{A+zz^*}(t) &= \det((tI - A) - zz^*) = \det(tI - A) \cdot (1 - z^*(tI - A)^{-1}z) \\ &= \left( \prod_{i \in [n]} (t - \lambda_i) \right) \left( 1 - \sum_{j \in [n]} \frac{|\langle v_j, z \rangle|^2}{t - \lambda_j} \right) \\ &= \prod_{i \in [n]} (t - \lambda_i) - \sum_{j \in [n]} |\langle v_j, z \rangle|^2 \prod_{\substack{i \in [n] \\ i \neq j}} (t - \lambda_j). \end{aligned}$$

We rewrite this expression in terms of the multiplicity  $m(\lambda')$  of each eigenvalue  $\lambda'$  of  $A$  to find polynomials  $q(t)$  and  $r(t)$  with  $q(\lambda) \neq 0 \neq r(\lambda)$  such that

$$\begin{aligned} p_{A+zz^*}(t) &= \prod_{\lambda'} (t - \lambda')^{m(\lambda')} - \sum_{\lambda'} \|P_{\lambda'} z\|^2 (t - \lambda')^{m(\lambda')-1} \prod_{\lambda'' \neq \lambda'} (t - \lambda'')^{m(\lambda'')} \\ &= (t - \lambda)^m q(t) - \|P_\lambda z\|^2 (t - \lambda)^{m-1} r(t). \end{aligned}$$

Since  $\|P_\lambda z\|^2 \neq 0$  by assumption, the result follows.  $\square$

**Lemma 6** *The preimage  $\Sigma^{-1}(\text{int}(E(d, n)))$  is dense in  $S(d, n)$ .*

**Proof** Given arbitrary  $Z = [z_1 \cdots z_n] \in S(d, n)$  and  $\epsilon > 0$ , we identify a matrix  $\tilde{Z} = [\tilde{z}_1 \cdots \tilde{z}_n] \in \Sigma^{-1}(\text{int}(E(d, n)))$  such that  $\|\tilde{Z} - Z\|_F < \epsilon$ . Put  $\tilde{z}_1 := z_1$ , and then for each  $j \in [n-1]$ , select any unit vector  $\tilde{z}_{j+1}$  in the open 2-ball of radius  $\epsilon/n$  about  $z_{j+1}$  that avoids the orthogonal complements of the eigenspaces of  $\sum_{i \in [j]} \tilde{z}_i \tilde{z}_i^*$ . Then  $\Sigma(\tilde{Z}) \in \text{int}(E(d, n))$  by Lemma 5 and  $\|\tilde{Z} - Z\|_F < \epsilon$  by the triangle inequality.  $\square$

Let  $\mathbb{T}^k \times \mathbb{U}(d-k)$  denote the group of matrices in  $\mathbb{U}(d)$  of the form

$$\begin{bmatrix} D & 0 \\ 0 & U \end{bmatrix}$$

with diagonal  $D \in \mathbb{U}(k)$ , and consider the following group of  $n$ -tuples of matrices:

$$\mathbb{U}(d, n) := \mathbb{U}(d) \times (\mathbb{T}^1 \times \mathbb{U}(d-1)) \times \cdots \times (\mathbb{T}^{d-1} \times \mathbb{U}(1)) \times (\mathbb{T}^d)^{n-d}.$$

(We use  $\mathbb{U}$  since each component of the tuple is *unitary*.) The following lemma describes how this set allows us to map back from  $\mathbb{E}(d, n)$  to  $\mathbb{S}(d, n)$ . This portion of our proof appears to break in the real case since the orthogonal group  $\mathbb{O}(1)$  is discrete; it would be interesting if the real case could be treated differently.

**Lemma 7** *There exists a continuous map  $\Phi: \mathbb{E}(d, n) \times \mathbb{U}(d, n) \rightarrow \mathbb{S}(d, n)$  such that*

- (a) *for every  $Z \in \mathbb{S}(d, n)$ , there exists  $U \in \mathbb{U}(d, n)$  such that  $\Phi(\Sigma(Z), U) = Z$ , and*
- (b) *for every  $\Lambda \in \mathbb{E}(d, n)$  and  $U \in \mathbb{U}(d, n)$ , it holds that  $\Sigma(\Phi(\Lambda, U)) = \Lambda$ .*

**Proof** We will define  $\Phi$  explicitly on  $\text{int}(\mathbb{E}(d, n)) \times \mathbb{U}(d, n)$  and then extend continuously to the full domain. For each  $\Lambda = [\lambda_1 \cdots \lambda_n] \in \text{int}(\mathbb{E}(d, n))$  and  $j \in [n]$ , denote  $r_j := \min\{j+1, d\}$ , define  $v_j, w_j \in \mathbb{R}^{r_j}$  by

$$(v_j)_i := \sqrt{\frac{\prod_{i' \in [r_j]} ((\lambda_j)_i - (\lambda_{j+1})_{i'})}{\prod_{\substack{i' \in [r_j] \\ i' \neq i}} ((\lambda_j)_i - (\lambda_j)_{i'})}}, \quad (w_j)_i := \sqrt{\frac{\prod_{i' \in [r_j]} ((\lambda_{j+1})_i - (\lambda_j)_{i'})}{\prod_{\substack{i' \in [r_j] \\ i' \neq i}} ((\lambda_{j+1})_i - (\lambda_{j+1})_{i'})}},$$

and define  $W_j \in \mathbb{U}(r_j)$  by

$$(W_j)_{i, i'} := \frac{(v_j)_i (w_j)_{i'}}{(\lambda_{j+1})_{i'} - (\lambda_j)_i}.$$

By Theorem 7 in [7],  $\Sigma^{-1}(\Lambda)$  equals the set of matrices  $[z_1 \cdots z_n] \in \mathbb{C}^{d \times n}$  obtained from all  $(U_1, V_1, \dots, V_{n-1}) \in \mathbb{U}(d, n)$  by taking  $z_1 := U_1 e_1$  and applying the following iteration for  $j \in [n-1]$ :

$$z_{j+1} := U_j V_j \begin{bmatrix} v_j \\ 0 \end{bmatrix}, \quad U_{j+1} := U_j V_j \begin{bmatrix} W_j & 0 \\ 0 & I_{d-r_j} \end{bmatrix}.$$

This process defines a map  $\Phi_0: \text{int}(\mathbb{E}(d, n)) \times \mathbb{U}(d, n) \rightarrow \mathbb{S}(d, n)$  such that

- (a') *for every  $Z \in \Sigma^{-1}(\text{int}(\mathbb{E}(d, n)))$ , there exists  $U \in \mathbb{U}(d, n)$  such that  $\Phi_0(\Sigma(Z), U) = Z$ , and*
- (b') *for every  $\Lambda \in \text{int}(\mathbb{E}(d, n))$  and  $U \in \mathbb{U}(d, n)$ , it holds that  $\Sigma(\Phi_0(\Lambda, U)) = \Lambda$ .*

Observe that  $R: \Lambda \mapsto (\{(v_j)_i^2\}_{i,j}, \{(W_j)_{i,i'}^2\}_{j,i,i'})$  is a rational map with bounded domain, and furthermore,  $R$  is bounded since  $\Phi_0$  is bounded. It follows that  $R$  is

Lipschitz, and so  $\Phi_0$  is uniformly continuous since  $\Phi_0(\Lambda, U)$  is a polynomial in  $(\sqrt{R(\Lambda)}, U)$  with bounded domain. In addition,  $S(d, n)$  is a complete metric space, and so there exists a unique continuous extension  $\Phi$  of  $\Phi_0$  to all of  $E(d, n) \times U(d, n)$ .

We now verify properties (a) and (b). For (a), take any  $Z \in S(d, n)$ . By Lemma 6, there exists a sequence  $\{Z_i\}_{i=1}^\infty$  in  $S(d, n)$  such that  $\Sigma(Z_i) \in \text{int}(E(d, n))$  for every  $i$  and  $Z_i \rightarrow Z$ . By (a'), it holds that for each  $i$ , there exists  $U_i \in U(d, n)$  such that  $\Phi(\Sigma(Z_i), U_i) = Z_i$ . By compactness, there is a subsequence of  $\{U_i\}_{i=1}^\infty$  with some limit point  $U \in U(d, n)$ . By continuity, it follows that  $\Phi(\Sigma(Z), U) = Z$ . For (b), take any  $\Lambda \in E(d, n)$  and  $U \in U(d, n)$ . There exists a sequence  $\{\Lambda_i\}_{i=1}^\infty$  in  $\text{int}(E(d, n))$  such that  $\Lambda_i \rightarrow \Lambda$ . By (b'), it holds that  $\Sigma(\Phi(\Lambda_i, U)) = \Lambda_i$  for each  $i$ . Since  $\Sigma$  and  $\Phi$  are continuous, taking limits of both sides gives  $\Sigma(\Phi(\Lambda, U)) = \Lambda$ .  $\square$

**Lemma 8** *The eigensteps map  $\Sigma: S(d, n) \rightarrow E(d, n)$  is open in the standard subspace topologies for  $S(d, n) \subseteq \mathbb{C}^{d \times n}$  and  $E(d, n) \subseteq \mathbb{R}^{d \times n}$ .*

**Proof** Take any  $Z \in S(d, n)$ . By Lemma 7(a), there exists  $U \in U(d, n)$  such that  $\Phi(\Sigma(Z), U) = Z$ . We may apply Lemma 4 with

$$f := \Sigma, \quad W := E(d, n), \quad g: \Lambda \mapsto \Phi(\Lambda, U).$$

Indeed, Lemma 7(a) gives  $(g \circ f)(Z) = \Phi(\Sigma(Z), U) = Z$  and Lemma 7(b) gives  $(f \circ g)(\Lambda) = \Sigma(\Phi(\Lambda, U)) = \Lambda$  for every  $\Lambda \in E(d, n)$ . Thus,  $\Sigma$  is open at  $Z$ . Since  $Z$  was arbitrary, it follows that  $\Sigma$  is open.  $\square$

We may now prove the main result of this section:

**Proof (of Theorem 2)** By Lemma 8, the eigensteps map  $\Sigma$  is open. As such,  $\text{FP}_f$  is the composition of the open map  $Z \mapsto \Sigma(Z)$  with the convex map  $\Lambda \mapsto f(\Lambda e_d)$ , and so the result follows from Lemma 3.  $\square$

## 4 Geometric invariant theory

The recent literature showcases an industry of using geometric methods to prove fundamental results in frame theory [26, 25, 27]. This section represents another step in this direction. We say  $Z \in \mathbb{C}^{d \times n}$  is **full spark** if every  $d \times d$  submatrix of  $Z$  is invertible [2]. We will prove the following strengthening of the Benedetto–Fickus theorem in the special case where  $n > d$ :

**Theorem 3** *Fix  $d, n \in \mathbb{N}$  with  $n > d$ , and consider the gradient flow  $F: S(d, n) \times [0, \infty) \rightarrow S(d, n)$  defined by the differential equation*

$$F(Z_0, 0) = Z_0, \quad \frac{d}{dt}F(Z_0, t) = -\text{grad FP}(F(Z_0, t)),$$

where  $\text{grad}$  denotes Riemannian gradient. If  $Z_0$  is full spark, then  $\lim_{t \rightarrow \infty} F(Z_0, t)$  is a unit norm tight frame.

Given a critical point that is not a unit norm tight frame, almost every member of any neighborhood of the critical point is full spark, which by Theorem 3 is sent by gradient flow to a unit norm tight frame. The absence of a basin of attraction indicates that the critical point is not a local minimizer. Since the frame potential is real analytic (in fact, polynomial), this can be made rigorous with a standard application of the Łojasiewicz inequality (à la [1]), thereby proving the Benedetto–Fickus theorem. Furthermore, Theorem 3 easily extends to the real case since the real submanifold of  $S(d, n)$  is invariant under gradient flow. To prove Theorem 3, we identify a property P such that

- every full spark member of  $S(d, n)$  satisfies P, and
- gradient flow (and its limit) preserves P, but
- no non-minimizing critical point of the frame potential satisfies P.

It turns out that this property originated from the *geometric invariant theory* developed by Mumford [24] using ideas from Hilbert [19]; see [34] for a survey. We start by introducing a few terms in some level of generality.

Let  $G$  be a reductive group acting linearly on a finite-dimensional complex vector space  $V$ , such as  $GL(V)$ ,  $SL(V)$ , or  $O(V)$ . A nonzero vector  $v \in V$  is said to be **unstable** under the action of  $G$  if the closure  $\text{cl}(G \cdot v)$  of the orbit of  $v$  contains 0, and otherwise  $v$  is **semi-stable** under the action of  $G$ .

**Proposition 6** *Given a nonzero vector  $v \in V$  that is semi-stable under the action of  $G$ , it holds that every member of  $\text{cl}(G \cdot v)$  is also semi-stable under the action of  $G$ .*

**Proof** We prove the contrapositive. Suppose there exists a nonzero  $w \in \text{cl}(G \cdot v)$  that is unstable under the action of  $G$ . Then there exist sequences  $\{g_i\}_{i=1}^{\infty}$  and  $\{h_j\}_{j=1}^{\infty}$  in  $G$  such that  $g_i \cdot v \rightarrow w$  and  $h_j \cdot w \rightarrow 0$ . For each  $j$ , the continuity of the linear action of  $h_j$  on  $V$  implies that

$$h_j \cdot w = h_j \cdot \lim_{i \rightarrow \infty} (g_i \cdot v) = \lim_{i \rightarrow \infty} (h_j g_i \cdot v) \in \text{cl}(G \cdot v),$$

and so  $0 = \lim_{j \rightarrow \infty} (h_j \cdot w) \in \text{cl}(G \cdot v)$ . Thus,  $v$  is unstable under the action of  $G$ .  $\square$

We are interested in the special case where  $G = SL(d)$  acts on  $V = (\mathbb{C}^d)^{\otimes n}$  by

$$g \cdot (z_1 \otimes \cdots \otimes z_n) := g z_1 \otimes \cdots \otimes g z_n$$

and extending linearly. We pass to this space by “tensoring up columns” with  $\tau: \mathbb{C}^{d \times n} \rightarrow (\mathbb{C}^d)^{\otimes n}$  defined by  $[z_1 \cdots z_n] \mapsto z_1 \otimes \cdots \otimes z_n$ . Then  $Z \in S(d, n)$  satisfies the previously mentioned property P if  $\tau(Z)$  is semi-stable under the action of  $SL(d)$ .

**Lemma 9** *Fix  $d, n \in \mathbb{N}$  with  $n > d$ .*

- (a) *For every full spark  $Z \in S(d, n)$ , it holds that  $\tau(Z)$  is semi-stable under the action of  $SL(d)$ .*
- (b) *For every non-minimizing critical point  $Z \in S(d, n)$  of the frame potential, it holds that  $\tau(Z)$  is unstable under the action of  $SL(d)$ .*

Assuming the lemma for now, we proceed with the proof of Theorem 3.

**Proof (sketch of Theorem 3)** For every  $Z \in S(d, n)$ , Lemma 1 gives the Euclidean gradient  $\nabla \text{EP}(Z) = 4ZZ^*Z$ , and so the Riemannian gradient  $\text{grad FP}(Z)$  is the matrix obtained by normalizing the columns of  $ZZ^*Z$ . Consider the action of  $\text{GL}(d) \times (\mathbb{C}^\times)^n$  on  $\mathbb{C}^{d \times n}$  defined by  $(g, \{a_i\}_{i \in [n]}) \cdot X := gX \text{diag}\{a_i\}_{i \in [n]}^{-1}$ . Then  $\text{grad FP}(Z)$  is in the tangent space at  $Z$  of the orbit  $(\text{GL}(d) \times (\mathbb{C}^\times)^n) \cdot Z$ , and so  $F(Z_0, t)$  resides in one such orbit for all  $t \geq 0$ . For every  $g \in \text{GL}(d)$ , it holds that  $g/\det(g)^{1/d} \in \text{SL}(d)$ , and so normalizing gives

$$F(Z_0, t) \in (\text{SL}(d) \times (\mathbb{C}^\times)^n) \cdot Z_0 \quad \forall t \geq 0.$$

Next, consider the action of  $\text{SL}(d) \times \mathbb{C}^\times$  on  $(\mathbb{C}^d)^{\otimes n}$  defined by  $(g, a) \cdot (z_1 \otimes \cdots \otimes z_n) := a(gz_1 \otimes \cdots \otimes gz_n)$  and extending linearly. Then

$$\tau(F(Z_0, t)) \in (\text{SL}(d) \times \mathbb{C}^\times) \cdot \tau(Z_0) \quad \forall t \geq 0.$$

Suppose  $Z_0$  is full spark. Then by Lemma 9(a),  $\tau(Z_0)$  is semi-stable under the action of  $\text{SL}(d)$ . Proposition 6 then implies that every point in  $\text{cl}((\text{SL}(d) \times \mathbb{C}^\times) \cdot \tau(Z_0))$  is also semi-stable under the action of  $\text{SL}(d)$ . Importantly, this includes  $\tau(Z_\infty)$ , where  $Z_\infty := \lim_{t \rightarrow \infty} F(Z_0, t)$ . As the limit point of gradient flow,  $Z_\infty$  is a critical point of the frame potential, and by Lemma 9(b),  $Z_\infty$  is a minimizer. It follows that  $Z_\infty$  is a unit norm tight frame.  $\square$

It remains to prove Lemma 9, which will require some technology from geometric invariant theory. For the moment, let's return to the general setting of a reductive group  $G$  acting linearly on a finite-dimensional complex vector space  $V$ . To discern whether a point is unstable under the action of  $G$ , it is convenient to work with small subgroups of  $G$ . A **one-parameter subgroup** of  $G$  is a homomorphism of algebraic groups  $\lambda: \mathbb{C}^\times \rightarrow G$ , meaning there exists a decomposition  $V = \bigoplus_{i \in I} V_i$  and integer weights  $w: I \rightarrow \mathbb{Z}$  such that for every  $i \in I$ ,  $v \in V_i$ , and  $t \in \mathbb{C}^\times$ , it holds that

$$\lambda(t) \cdot v = t^{w(i)} v.$$

Of course, we may conclude that  $v \in V \setminus \{0\}$  is unstable under the action of  $G$  if there exists a one-parameter subgroup  $\lambda$  of  $G$  such that

$$\lim_{t \rightarrow 0} \lambda(t) \cdot v = 0.$$

Amazingly, the converse also holds:

**Proposition 7 (Hilbert–Mumford criterion [19, 24])** *A nonzero vector  $v \in V$  is unstable under the action of  $G$  if and only if there exists a one-parameter subgroup  $\lambda$  of  $G$  such that*

$$\lim_{t \rightarrow 0} \lambda(t) \cdot v = 0.$$

**Proof (of Lemma 9)** (a) Select any one-parameter subgroup  $\lambda: \mathbb{C}^\times \rightarrow \text{SL}(d)$ . By the determinant, the integer weights  $w: [d] \rightarrow \mathbb{Z}$  of  $\lambda$  satisfy  $\sum_{i \in [d]} w(i) = 0$ .



Reindex as necessary so that  $w(1) \geq \dots \geq w(d)$ . If  $w(i) = 0$  for every  $i \in [d]$ , then  $\lambda(t)$  equals the identity matrix for all  $t \in \mathbb{C}^\times$ , and so  $\lim_{t \rightarrow 0} \lambda(t) \cdot \tau(Z) = \tau(Z) \neq 0$ . Otherwise,  $w(1) > 0 > w(d)$ . Put  $D(t) := \text{diag}\{t^{w(i)}\}_{i \in [d]}$  and take  $S \in \text{GL}(d)$  such that  $\lambda(t) = SD(t)S^{-1}$ . Given a full spark  $Z \in \text{S}(d, n)$ , put  $\tilde{Z} := [\tilde{z}_1 \cdots \tilde{z}_n] := S^{-1}Z$ . Then

$$\begin{aligned} S^{-1}\lambda(t) \cdot \tau(Z) &= D(t)S^{-1} \cdot (z_1 \otimes \cdots \otimes z_n) \\ &= D(t)\tilde{z}_1 \otimes \cdots \otimes D(t)\tilde{z}_n \\ &= \left( \sum_{i_1 \in [d]} t^{w(i_1)} (\tilde{z}_1)_{i_1} e_{i_1} \right) \otimes \cdots \otimes \left( \sum_{i_n \in [d]} t^{w(i_n)} (\tilde{z}_n)_{i_n} e_{i_n} \right) \\ &= \sum_{i_1, \dots, i_n \in [d]} \left( t^{\sum_{j \in [n]} w(i_j)} \prod_{j \in [n]} (\tilde{z}_j)_{i_j} \right) e_{i_1} \otimes \cdots \otimes e_{i_n}. \end{aligned} \quad (5)$$

We will use the fact that  $Z$  (and thus  $\tilde{Z}$ ) is full spark to identify a component of (5) that does not vanish as  $t \rightarrow 0$ . By the Hilbert–Mumford criterion, this suffices to prove the result.

Let  $k$  denote the number of entries in the  $d$ th row of  $\tilde{Z}$  that equal zero, let  $j_1, \dots, j_k \in [n]$  denote the corresponding column indices, and index the remaining indices as  $\{j_{k+1}, \dots, j_n\} := [n] \setminus \{j_1, \dots, j_k\}$ . (In the degenerate case where  $k = 0$ , we may take  $j_\ell = \ell$  for each  $\ell \in [n]$ .) Writing  $\tilde{Z} = [\tilde{z}_1 \cdots \tilde{z}_n]$ , then full spark implies

$$0 \neq \det([\tilde{z}_{j_1} \cdots \tilde{z}_{j_d}]) = \sum_{\sigma \in S_d} \text{sign}(\sigma) \prod_{t \in [d]} (\tilde{z}_{j_t})_{\sigma(t)}.$$

It follows that there exists a permutation  $\sigma \in S_d$  such that

$$\prod_{t \in [d]} (\tilde{z}_{j_t})_{\sigma(t)} \neq 0.$$

In addition,  $(\tilde{z}_{j_t})_d \neq 0$  for every  $t > k$  by construction, and full spark implies  $k < d$ . As such, we consider the component of (5) with indices

$$i_j := \begin{cases} \sigma(t) & \text{if } j = j_t, t \in [d] \\ d & \text{otherwise} \end{cases}$$

since

$$\prod_{j \in [n]} (\tilde{z}_j)_{i_j} = \left( \prod_{t \in [d]} (\tilde{z}_{j_t})_{\sigma(t)} \right) \left( \prod_{t \in \{d+1, \dots, n\}} (\tilde{z}_{j_t})_d \right) \neq 0$$

and

$$\sum_{j \in [n]} w(i_j) = \sum_{i \in [d]} w(i) + (n-d)w(d) = (n-d)w(d) < 0,$$

where the last inequality uses the assumption that  $d < n$  and the fact that  $w(d) < 0$ . It follows that the  $(i_1, \dots, i_n)$  component of (5) diverges as  $t \rightarrow 0$ .

(b) Suppose  $Z \in S(d, n)$  is a non-minimizing critical point of the frame potential. By Corollary 1, there exist  $d', n' \in \mathbb{N}$  satisfying

$$\frac{n'}{d'} > \frac{n}{d} \quad (6)$$

such that  $n'$  of the columns of  $Z$  form a tight frame for a  $d'$ -dimensional subspace of  $\mathbb{C}^d$  that is orthogonal to the other  $n - n'$  columns. Explicitly, there is a unitary matrix  $U \in \mathbb{C}^{d \times d}$  and a permutation matrix  $P \in \mathbb{R}^{n \times n}$  such that

$$UZP^* = \begin{bmatrix} X & 0 \\ 0 & Y \end{bmatrix},$$

where  $X \in \mathbb{C}^{d' \times n'}$  satisfies  $XX^* = \frac{n'}{d'} I_{d'}$ . Put

$$D(t) := \begin{bmatrix} t^{d-d'} I_{d'} & 0 \\ 0 & t^{-d'} I_{d-d'} \end{bmatrix},$$

and take the one-parameter subgroup  $\lambda: \mathbb{C}^\times \rightarrow \mathrm{SL}(d)$  defined by  $\lambda(t) := U^{-1} D(t) U$ . Denoting  $Z = [z_1 \cdots z_n]$  and  $UZ = [\tilde{z}_1 \cdots \tilde{z}_n]$ , then

$$\begin{aligned} U\lambda(t) \cdot \tau(Z) &= D(t)U \cdot (z_1 \otimes \cdots \otimes z_n) \\ &= D(t)\tilde{z}_1 \otimes \cdots \otimes D(t)\tilde{z}_n = (t^{d-d'})^{n'} (t^{-d'})^{n-n'} \tilde{z}_1 \otimes \cdots \otimes \tilde{z}_n. \end{aligned}$$

The exponent on  $t$  simplifies to  $n'd - d'n$ , which is positive by (6). This implies that  $\lim_{t \rightarrow 0} \lambda(t) \cdot \tau(Z) = 0$ , as claimed.  $\square$

To conclude, we note that in Section 16.3 of the monograph [21], Kirwan reports that  $\tau(Z)$  is semi-stable under the action of  $\mathrm{SL}(d)$  precisely when

$$\frac{|\{j : z_j \in S\}|}{\dim(S)} \leq \frac{n}{d}$$

for every nontrivial subspace  $S$  of  $\mathbb{C}^d$ . Interestingly, the identical condition appears in Barthe's theorem [18] to characterize the real  $Z \in S(d, n)$  for which there exists  $A \in \mathrm{GL}(d)$  such that normalizing the columns of  $AZ$  gives a unit norm tight frame. While the original proof of Barthe's theorem seems difficult to parse [3], this coincidence suggests a geometric proof of the complex case using the fact that gradient flow sends semi-stable points to unit norm tight frames.

**Acknowledgements** Some of the ideas in this project were conceived at the 2018 MFO Mini-Workshop on Algebraic, Geometric, and Combinatorial Methods in Frame Theory. DGM was partially supported by AFOSR FA9550-18-1-0107 and NSF DMS 1829955. TN was partially supported by NSF DMS 2107808. SV was partially supported by NSF DMS 2044349, the NSF–Simons Research Collaboration on the Mathematical and Scientific Foundations of Deep Learning (MoDL) (NSF DMS 2031985), and the TRIPODS Institute for the Foundations of Graph and Deep Learning at Johns Hopkins University. CS was partially supported by NSF DMS 2107700.

## References

1. P.-A. Absil, K. Kurdyka, On the stable equilibrium points of gradient systems, *Syst. Control. Lett.* 55 (2006) 573–577.
2. B. Alexeev, J. Cahill, D. G. Mixon, Full spark frames, *J. Fourier Anal. Appl.* 18 (2012) 1167–1194.
3. F. Barthe, On a reverse form of the Brascamp–Lieb inequality, *Invent. Math.* 134 (1998) 335–361.
4. J. J. Benedetto, M. Fickus, Finite normalized tight frames, *Adv. Comput. Math.* 18 (2003) 357–385.
5. D. P. Bertsekas, *Nonlinear Programming*, 2nd ed., Athena Scientific, 1999.
6. N. Boumal, V. Voroninski, A. Bandeira, The non-convex Burer–Monteiro approach works on smooth semidefinite programs, *Adv. Neural Inf. Process. Syst.* 29 (2016) 2757–2765.
7. J. Cahill, M. Fickus, D. G. Mixon, M. J. Poteet, N. Strawn, Constructing finite frames of a given spectrum and set of lengths, *Appl. Comput. Harmon. Anal.* 35 (2013) 52–73.
8. J. Cahill, D. G. Mixon, N. Strawn, Connectivity and irreducibility of algebraic varieties of finite unit norm tight frames, *SIAM J. Appl. Algebra Geom.* 1 (2017) 38–72.
9. E. J. Candès, X. Li, M. Soltanolkotabi, Phase retrieval via Wirtinger flow: Theory and algorithms, *IEEE Trans. Inform. Theory* 61 (2015) 1985–2007.
10. P. G. Casazza, J. Kovačević, Equal-norm tight frames with erasures, *Adv. Comput. Math.* 18 (2003) 387–430.
11. I. Daubechies, A. Grossmann, Y. Meyer, Painless nonorthogonal expansions, *J. Math. Phys.* 27 (1986) 1271–1283.
12. J. A. De Loera, T. B. McAllister, Vertices of Gelfand–Tsetlin polytopes, *Discrete Comput. Geom.* 32 (2004) 459–470.
13. R. J. Duffin, A. C. Schaeffer, A class of nonharmonic Fourier series, *Trans. Amer. Math. Soc.* 72 (1952) 341–366.
14. M. Fickus, D. G. Mixon, Tables of the existence of equiangular tight frames, arXiv:1504.00253.
15. R. Ge, F. Huang, C. Jin, Y. Yuan, Escaping from saddle points—online stochastic gradient for tensor decomposition, *Proc. Mach. Learn. Res.* 40 (2015) 797–842.
16. R. Ge, J. D. Lee, T. Ma, Matrix completion has no spurious local minimum, *Adv. Neural Inf. Process. Syst.* 29 (2016) 2981–2989.
17. V. K. Goyal, J. Kovačević, J. A. Kelner, Quantized frame expansions with erasures, *Appl. Comput. Harmon. Anal.* 10 (2001) 203–233.
18. L. Hamilton, A. Moitra, The Paulsen problem made simple, *Isr. J. Math.*, to appear.
19. D. Hilbert, Über die vollen Invariantensysteme, *Math. Annalen*, 42 (1893) 313–373.
20. R. B. Holmes, V. I. Paulsen, Optimal frames for erasures, *Linear Algebra Appl.* 377 (2004) 31–51.
21. F. C. Kirwan, *Cohomology of Quotients in Symplectic and Algebraic Geometry*, Mathematical Notes, vol. 31, Princeton U. Press, Princeton, NJ, 1984.
22. K. Kreutz-Delgado, The Complex Gradient Operator and the  $\mathbb{C}\mathbb{R}$ -Calculus, arXiv:0906.4835.
23. T. Laurent, J. Brecht, Deep linear networks with arbitrary loss: All local minima are global, *Proc. Mach. Learn. Res.* 80 (2018) 2902–2907.
24. D. Mumford, J. Fogarty, F. Kirwan, *Geometric Invariant Theory*, Springer Science & Business Media, 1994.
25. T. Needham, C. Shonkwiler, Admissibility and frame homotopy for quaternionic frames, arXiv:2108.02275.
26. T. Needham, C. Shonkwiler, Symplectic geometry and connectivity of spaces of frames, *Adv. Comput. Math.* 47 (2021) 5.
27. T. Needham, C. Shonkwiler, Toric symplectic geometry and full spark frames, arXiv:2110.11295.
28. M. Nouiehed, M. Razaviyayn, Learning deep models: Critical points and local openness, ICLR 2018, arXiv:1803.02968.
29. S. Smale, Mathematical problems for the next century, *Math. Intell.* 20 (1998) 7–15.

30. T. Strohmer, R. W. Heath, Grassmannian frames with applications to coding and communication, *Appl. Comput. Harmon. Anal.* 14 (2003) 257–275.
31. J. Sun, Q. Qu, J. Wright, Complete dictionary recovery over the sphere I: Overview and the geometric picture, *IEEE Trans. Inform. Theory* 63 (2016) 853–884.
32. J. Sun, Q. Qu, J. Wright, A geometric analysis of phase retrieval, *Found. Comput. Math.* 18 (2018) 1131–1198.
33. P. M. L. Tammes, On the origin of number and arrangement of the places of exit on the surface of pollen-grains, *Recueil Trav. Bot. Néerl.* 27 (1930) 1–84.
34. R. P. Thomas, Notes on GIT and symplectic reduction for bundles and varieties, in: *Surveys in Differential Geometry, Vol. X: Essays in Geometry in Memory of S. S. Chern*, International Press, 2006, pp. 221–273.
35. J. J. Thomson, On the structure of the atom, *Phil. Mag.* 7 (1904) 237–265.
36. L. Welch, Lower bounds on the maximum cross correlation of signals, *IEEE Trans. Inform. Theory* 20 (1974) 397–399.