Frames, the Loewner order and eigendecomposition for morphological operators on tensor fields

Jasper J. van de Gronde, Jos B.T.M. Roerdink

Johann Bernoulli Institute for Mathematics and Computer Science,
University of Groningen, P.O. Box 407, 9700 AK Groningen, The Netherlands

Abstract

Rotation invariance is an important property for operators on tensor fields, but up to now, most methods for morphology on tensor fields had to either sacrifice rotation invariance, or do without the foundation of mathematical morphology: a lattice structure. Recently, we proposed a framework for rotation-invariant mathematical morphology on tensor fields that does use a lattice structure. In addition, this framework can be derived systematically from very basic principles. Here we show how older methods for morphology on tensor fields can be interpreted within our framework. On the one hand this improves the theoretical underpinnings of these older methods, and on the other this opens up possibilities for improving the performance of our method. We discuss commonalities and differences of our method and two methods developed by Burgeth et al.

Keywords: mathematical morphology, tensor fields, frames, rotation invariance, Loewner order

1. Introduction

Although the theory of mathematical morphology is well developed for scalar images, the generalization to tensor-valued images is not straightforward. With increasing interest in processing tensor fields like flow data and diffusion MRI scans, it is becoming increasingly important to solve this problem. Recently, we suggested a method for generalizing morphological operators defined on scalar

images to operators on vector-valued images [2, 3] and tensor fields [4], using group invariance to avoid some of the pitfalls one encounters with naive generalizations.

There are currently several ways of approaching tensor morphology. First of all, one can consider the (scalar) differences between neighbouring tensors [5, 6]. Another approach uses a preorder on tensors derived from the data [7]. These are valid and interesting approaches, but not direct generalizations of the traditional lattice framework on greyscale images. One can also construct total orders on tensors in various ways [8]. However, total orders on higher dimensional spaces lead to discontinuous supremum and infimum (join and meet) operators.

Also interesting are some approaches where existing formulas for the supremum/infimum of a set of real numbers are generalized to tensors [8, 9]. So far, however, little is known about the properties of the resulting operators and how they compare to other approaches.

Some more direct generalizations of traditional morphological theory are given by [1, 9–11]. Having started out with a partial order on matrices based on set inclusion of the image of a matrix acting on a unit ball, these authors have ultimately settled on using a partial order on (symmetric) matrices called the Loewner order. Unfortunately the Loewner order does not give rise to a so-called "lattice", and is therefore not directly interpretable in the traditional morphological framework.

Our frame-based method works by lifting the tensor field to a higher dimensional representation in which it is much easier to define morphological operators that are rotation invariant. This representation is effectively equivalent to the representation used by Duits et al. [12]. However, both methods arrive at this representation through different means. Also, Duits et al. focus on what happens after you have this representation, developing a particular method for computing dilations and erosions. In our case, we have also provided a way of going back to the original tensor field representation with minimal loss in a least-squares sense.

Here we focus on unifying our approach with Burgeth et al.'s approach.

We show that so far the two main approaches for a direct generalization of mathematical morphology to tensor-valued images are deeply connected with each other. In doing so, we open the door to further experimentation along the same lines, and gain insight into what trade-offs are involved.

In particular, we show that the Loewner-order-based methods developed by Burgeth et al. [1] can be interpreted as implicitly using our frame-based method with a different projection back to the original space. We also prove a similar result for the approach used by Burgeth et al. to "lift" functions on scalars to functions on tensors [1, 13, 14]. This is then applied to Burgeth et al.'s approach for morphological filtering of tensor fields using partial differential equations.

2. Definitions and notation

2.1. Mathematical morphology

Shapes are central in mathematical morphology. Originally these shapes were connected components in binary images, and the theory was based on sets. The modern approach is to view an image as an element of a so-called lattice. A *lattice* is a partially ordered set for which every pair of elements has a unique least upper bound (usually called supremum or join) and a unique greatest lower bound (infimum or meet). The join is denoted by '∨' and the meet by '∧'. Many useful operators, like erosion and dilation, can be developed for lattices. An erosion is any operator that commutes with taking the meet, often constructed for images by taking the meet over a certain neighbourhood around each point. A dilation is like an erosion, but commuting with the join (for any dilation there is a corresponding erosion and vice versa).

If the lattice is also a vector space and we want the lattice to be compatible with the vector space structure, then it has been shown [15, thm. XV.1] that the lattice should be built using "direct and lexicographical union". Since lexicographical orders give rise to discontinuous joins and meets (such that an arbitrarily small change in the input can lead to an arbitrarily large change in the output), we will stick to "direct" union, which leads to a *product order*. This

3

75 means that vectors must essentially be ordered on the coefficients in some basis,

76 and that one vector is less than or equal to another if all of its coefficients are

77 less than or equal to those of the other. If $a_k \in \mathbb{R}$ is the coefficient of vector $\mathbf{a}$

78 with index $k \in \mathcal{K}$, then the join and meet are defined as follows: for all $k \in \mathcal{K}$

79
80
$$(\mathbf{a} \vee \mathbf{b})_k = a_k \vee b_k \qquad\qquad (\mathbf{a} \wedge \mathbf{b})_k = a_k \wedge b_k.$$

81 Clearly, the set of vectors greater than or equal to zero forms a cone: $\mathbf{a} \geq 0 \implies$

82 $\lambda \mathbf{a} \geq 0$ for all positive λ, and $\mathbf{a} \geq 0$ and $\mathbf{b} \geq 0 \implies \mathbf{a} + \mathbf{b} \geq 0$. This cone (C)

83 is called the *ordering cone*, as we have $\mathbf{a} \leq \mathbf{b} \iff \mathbf{b} - \mathbf{a} \in C$.

84 **Example 1.** *(Product order on vectors).* Suppose that we consider a two-

85 dimensional vector space with a product order based on the basis $\{\mathbf{e}_k\}_{k \in \{1,2\}}$.

86 Then, if we have two vectors $\mathbf{a} = \mathbf{e}_1$ and $\mathbf{b} = \mathbf{e}_2$, then their meet $\mathbf{a} \wedge \mathbf{b}$ would

87 equal 0, while their join would equal $\mathbf{e}_1 + \mathbf{e}_2$. Similarly, if $\mathbf{a}$ would equal $\mathbf{e}_1$ and

88 $\mathbf{b}$ would equal $\frac{1}{2}\mathbf{e}_1 + \mathbf{e}_2$, we have $\mathbf{a} \wedge \mathbf{b} = \frac{1}{2}\mathbf{e}_1$ and $\mathbf{a} \vee \mathbf{b} = \mathbf{e}_1 + \mathbf{e}_2$.

89 *2.2. Tensors*

90 We only consider real, symmetric tensors (sometimes called supersymmet-

91 ric), and look at them as being generated by the symmetrized tensor product

92 '$\odot$'. The vector space on which the tensors are based is denoted by V and is

93 taken to be finite-dimensional, while the space of degree-n symmetric tensors is

94 denoted by $V^{\odot n}$ (which itself is also a vector space). A tensor of degree n is

95 symmetric if and only if it can be written as a sum of tensor "powers" (see [16,

96 Lemma 4.2])

97
$$\mathbf{a}^{\odot n} = \underbrace{\mathbf{a} \odot \cdots \odot \mathbf{a}}_{n \text{ times}} = \underbrace{\mathbf{a} \otimes \cdots \otimes \mathbf{a}}_{n \text{ times}} = \mathbf{a}^{\otimes n}.$$

Here '$\otimes$' is the regular tensor product, and '$\odot$' is the symmetrized tensor prod-

uct as discussed by Kostrikin and Manin [17, ch. 4 Proposition 5.7][1]. The sym-

---

[1]In some cases a different symmetric tensor product is used in the literature, which (for
a vector space over the reals) is essentially the same up to multiplication by $n!$. Kostrikin
and Manin [17, ch. 4 §5.9] and Bourbaki [18, §III.6] both discuss this approach in some detail.
The symmetric (or symmetrized) tensor product we use is sometimes denoted without explicit
operator (so $\mathbf{a} \odot \mathbf{b}$ would be written as $\mathbf{ab}$). We consistently use the operator '$\odot$'.

4

metrized tensor product can be defined as giving the average of all permutations of the regular tensor product. So for example:

$$\mathbf{a} \odot \mathbf{b} \odot \mathbf{c} = \frac{1}{6}(\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c} + \mathbf{b} \otimes \mathbf{a} \otimes \mathbf{c}$$
$$+ \mathbf{a} \otimes \mathbf{c} \otimes \mathbf{b} + \mathbf{b} \otimes \mathbf{c} \otimes \mathbf{a}$$
$$+ \mathbf{c} \otimes \mathbf{a} \otimes \mathbf{b} + \mathbf{c} \otimes \mathbf{b} \otimes \mathbf{a}).$$

It will be assumed that V is not just a vector space, but that it is a Hilbert space with the inner product denoted by '$\cdot$'. The inner product on symmetric tensors can be defined based on its linearity and the inner product on the underlying vector space ($\mathbf{a}, \mathbf{b} \in V$):

$$\mathbf{a}^{\odot n} \cdot \mathbf{b}^{\odot n} = (\mathbf{a} \cdot \mathbf{b})^n.$$

Sometimes it is useful to make use of the fact that if $\mathbf{B}_1$ is a degree-m_1 tensor and $\mathbf{B}_2$ is a degree-m_2 tensor, such that $m_1 + m_2 = n$, then

$$\mathbf{a}^{\odot n} \cdot (\mathbf{B}_1 \odot \mathbf{B}_2) = (\mathbf{a}^{\odot m_1} \cdot \mathbf{B}_1)(\mathbf{a}^{\odot m_2} \cdot \mathbf{B}_2). \tag{1}$$

The symmetric degree-n identity tensor $\mathbf{I}_n$ is defined here[2] as the unique symmetric tensor that satisfies $\mathbf{a}^{\odot n} \cdot \mathbf{I}_n = \|\mathbf{a}\|^n$ for all non-zero $\mathbf{a} \in V$. It should be noted that this definition only makes sense for tensor spaces of even degree, as for odd n we would have $\mathbf{a}^{\odot n} \cdot \mathbf{I}_n = -(-\mathbf{a})^{\odot n} \cdot \mathbf{I}_n$. For even n it can be shown that such a tensor *does* exist, and is uniquely defined. First of all, if it exists it is obviously unique, as a symmetric tensor is characterized completely by its inner products with tensor powers of the form $\mathbf{a}^{\odot n}$. Secondly, we can see that it exists by using either of the following equivalent constructions:

$$\mathbf{I}_n = D_{n,V} \int_{S_V} \mathbf{s}^{\odot n} \, \mathrm{d}\mathbf{s} \qquad \text{or} \qquad \mathbf{I}_n = \mathbf{I}_2^{\odot n/2}.$$

---

[2]This convention for defining the identity tensor is not very common, but it is interesting to note that Qi [19] introduces it as I_E to allow the computation of rotation-invariant "eigenvalues" of higher degree tensors. Also, the most common alternative (a diagonal tensor with ones on the diagonal) seems to have little intrinsic meaning (in particular, it is not rotation invariant).

Here S_V is the sphere of all unit length vectors in V and $D_{n,V}$ is a (positive real) constant fully determined by $\mathbf{a}^{\odot n} \cdot \mathbf{I}_n = \|\mathbf{a}\|^n$ (for any non-zero $\mathbf{a} \in V$), while $\mathbf{I}_2$ can be seen to be $\sum_{k \in \mathcal{K}} \mathbf{e}_k^{\odot 2}$ for any orthonormal basis $\{\mathbf{e}_k\}_{k \in \mathcal{K}}$ of V. The degree-two identity tensor is typically represented by a matrix whose diagonal entries are all one, and whose off-diagonal entries are all zero. However, it should be noted that for n larger than two, $\mathbf{I}_n$ is typically *not* a "diagonal" tensor. That is, if one were to write it as an n-dimensional array (based on an orthogonal basis for the underlying vector space), it would contain off-diagonal values (see Example 2). In addition, if one considers not just symmetric tensors one could define asymmetric "identity tensors" (representing the identity map on degree-two tensors for example). Note that the integral construction can be seen to work from the point of view of symmetry (assuming a rotationally invariant measure on the unit sphere), while the construction based on $\mathbf{I}_2$ can be seen to work using Eq. (1).

In some cases it is useful to "lift" functions on the vector space V to the tensor space $V^{\odot n}$: given a (linear) function $R : V \to V$, the function $R^{\odot n} : V^{\odot n} \to V^{\odot n}$ is defined by $R^{\odot n}(\mathbf{a}^{\odot n}) = (R(\mathbf{a}))^{\odot n}$. For linear functions we often leave out the parentheses around the parameter, so we would have: $R^{\odot n} \mathbf{a}^{\odot n} = (R\mathbf{a})^{\odot n}$.

Example 2. *(Representing tensors as supermatrices).* In some cases (especially in implementations) it can be useful to represent tensors using arrays of numbers. As this kind of representation can be considered a generalization of representing degree-two tensors using matrices, the resulting arrays are sometimes called supermatrices.

The first step is to pick a basis $\{\mathbf{e}_k\}_{k \in \mathcal{K}}$, often assumed to be orthonormal. A degree-n (not necessarily symmetric) tensor $\mathbf{A}$ can then be represented as:

$$\mathbf{A} = \sum_{k \in \mathcal{K}^n} A_k \, (\mathbf{e}_{k_1} \otimes \cdots \otimes \mathbf{e}_{k_n}).$$

For a symmetric tensor, A_k should equal A_m if some permutation of m is equal to k. Alternatively, we can sum only over the k that are in non-descending order and use $\mathbf{e}_{k_1} \odot \cdots \odot e_{k_n}$ as the basis vectors.

We can use the above to show that higher degree identity tensors are not diagonal. Assuming an orthonormal basis $\{\mathbf{e}_k\}_{k\in\mathcal{K}}$,

$$
\begin{aligned}
\mathbf{I}_4 = \mathbf{I}_2^{\odot 2} &= \Big(\sum_{k\in\mathcal{K}} \mathbf{e}_k^{\odot 2}\Big) \odot \Big(\sum_{k\in\mathcal{K}} \mathbf{e}_k^{\odot 2}\Big) \\
&= \sum_{k\in\mathcal{K}}\sum_{m\in\mathcal{K}} \mathbf{e}_k^{\odot 2} \odot \mathbf{e}_m^{\odot 2} \\
&= \sum_{k\in\mathcal{K}}\sum_{m\in\mathcal{K}} \frac{1}{6}(\mathbf{e}_k \otimes \mathbf{e}_k \otimes \mathbf{e}_m \otimes \mathbf{e}_m + \mathbf{e}_k \otimes \mathbf{e}_m \otimes \mathbf{e}_k \otimes \mathbf{e}_m \\
&\qquad\qquad + \mathbf{e}_k \otimes \mathbf{e}_m \otimes \mathbf{e}_m \otimes \mathbf{e}_k + \mathbf{e}_m \otimes \mathbf{e}_k \otimes \mathbf{e}_k \otimes \mathbf{e}_m \\
&\qquad\qquad + \mathbf{e}_m \otimes \mathbf{e}_k \otimes \mathbf{e}_m \otimes \mathbf{e}_k + \mathbf{e}_m \otimes \mathbf{e}_m \otimes \mathbf{e}_k \otimes \mathbf{e}_k).
\end{aligned}
$$

Comparing this to the above, we see that the diagonal entries (corresponding to $k = m$) are equal to one, but that there are also some non-zero off-diagonal entries (wherever $k \neq m$). Note that each of these off-diagonal entries occurs twice in the sum (once for $k = k_1 \in \mathcal{K}$ and $m = k_2 \in \mathcal{K}$ and once for $k = k_2$ and $m = k_1$), so that all these off-diagonal entries equal one-third.

3. Summary of discussed methods

Here we give an overview of the different ways of approaching tensor morphology that we will consider.

3.1. Group-invariant frames

Tensors of degree n can be viewed to form a vector space. Rather than representing tensors in this vector space as linear combinations of basis vectors (tensors), we can also use a *frame*: a set of vectors (tensors) $\{\mathbf{F}_i\}_{i\in\mathcal{I}}$ (not necessarily finite or even countable) for which there are finite, positive constants A and B such that

$$
A \|\mathbf{A}\|^2 \leq \|F\mathbf{A}\|^2 \leq B \|\mathbf{A}\|^2,
$$

for any $\mathbf{A} \in V^{\odot n}$. Here the linear operator $F : V^{\odot n} \to \mathbb{R}^{\mathcal{I}}$ is called the analysis operator, and is defined by $(F\mathbf{A})_i = \mathbf{F}_i \cdot \mathbf{A}$ for all $i \in \mathcal{I}$. The pseudo-inverse F^+ of F can be used to go back from $\mathbb{R}^{\mathcal{I}}$ to $V^{\odot n}$ in a least-squares manner ([20,

7

§1.4] and [21, ch. 1 and 8]). We will use the rotation-invariant frame developed previously [4]. This frame corresponds to the set of tensors that can be written as $\mathbf{s}^{\odot n}$ with $\mathbf{s}$ any unit vector in V (so $\mathbf{s} \in S_V$). So instead of $\mathbb{R}^{\mathcal{I}}$ we will write $\mathbb{R}^{S_V}$, or $\mathbb{R}^S$ for short.

Although the above defines what vectors are in the frame we use, it does not discuss how the frame components can be indexed. In its raw form, our original construction suggested indexing by group actions [2]. However, in the current work we will simply index by vectors in the unit d-sphere (consistent with writing $\mathbb{R}^{S_V}$). This does not compromise the rotation invariance, and allows us to abstract from the tensor degree in some places. So throughout this paper, F_n will refer to the analysis operator of the frame consisting of all n-fold tensor powers of d-dimensional unit vectors, indexed by those unit vectors. In some cases the subscript will be dropped, if it is clear from the context. We thus have $(F_n \mathbf{A})_{\mathbf{s}} = \mathbf{s}^{\odot n} \cdot \mathbf{A}$ for all $\mathbf{s} \in S_V$ and $\mathbf{A} \in V^{\odot n}$.

Note that we use the convention that a function h on the reals can be applied to the frame representation of a tensor by acting on all coefficients individually. Similarly, the meet and join (infimum and supremum) act per-channel, leading to a product order on the frame coefficients. The inner product on the frame coefficients is defined as follows ($\mathbf{u}, \mathbf{v} \in \mathbb{R}^S$):

$$\mathbf{u} \cdot \mathbf{v} = \int_{S_V} u_{\mathbf{s}} v_{\mathbf{s}} \, \mathrm{d}\mathbf{s}.$$

Here there is no dependence on the degree of the original tensors, as all tensors, regardless of degree, map to the same space of frame coefficients: $\mathbb{R}^{S_V}$. This is a consequence of our choice to index the frame representation by unit vectors.

Note that in practice the easiest way to work with group-invariant frames is to take a finite, uniformly distributed, subset of the frame vectors. The analysis operator is then represented by a matrix F whose rows are the frame vectors. This is the approach used for the examples in this paper.

Example 3. *(The trace and the identity tensor).* For matrices, the trace is the sum of the diagonal entries. Assuming an orthonormal basis this can easily be seen to be equivalent to taking the inner product with the (degree-two) identity

tensor. In the frame representation, the identity tensor (of any degree) is lifted to $F_n \mathbf{I}_n = \mathbf{i} \in \mathbb{R}^S$ such that $\mathbf{i_s} = 1$ for all $\mathbf{s} \in S_V$. Interestingly, $\mathbf{A} \cdot \mathbf{I}_n$ and $F_n \mathbf{A} \cdot \mathbf{i}$ are closely related:

$$\begin{aligned}
\mathbf{A} \cdot \mathbf{I}_n = \mathbf{A} \cdot \left(D_{n,V} \int_{S_V} \mathbf{s}^{\odot n} \, \mathrm{d}\mathbf{s} \right) &= D_{n,V} \int_{S_V} \mathbf{A} \cdot \mathbf{s}^{\odot n} \, \mathrm{d}\mathbf{s} \\
&= D_{n,V} \int_{S_V} (F_n \mathbf{A})_{\mathbf{s}} \, \mathbf{i_s} \, \mathrm{d}\mathbf{s} = D_{n,V} \, (F_n \mathbf{A} \cdot \mathbf{i}).
\end{aligned}$$

This can be seen as (extra) motivation for generalizing the trace to higher degree tensors using the inner product with the identity tensor.

3.2. The Loewner order

The Loewner order on matrices considers a matrix $\mathbf{A}$ less than or equal to a matrix $\mathbf{B}$ if and only if $\mathbf{B} - \mathbf{A}$ is positive semidefinite; this relation is denoted by $\mathbf{A} \leq_L \mathbf{B}$. A matrix $\mathbf{A}$, here viewed as a degree-two tensor, is positive semidefinite if and only if $\mathbf{A} \cdot \mathbf{v}^{\odot 2} \geq 0$ for all non-zero $\mathbf{v} \in V$. By analogy, we will consider any degree-n tensor to be positive semidefinite if and only if $\mathbf{A} \cdot \mathbf{v}^{\odot n} \geq 0$ for all non-zero $\mathbf{v} \in V$. It should be clear from this definition that only tensors of even degree can be positive semidefinite.

It is interesting to examine the ordering cone connected with the Loewner order. Clearly this cone consists of all positive semidefinite tensors (of degree n). All tensors of the form $\mathbf{a}^{\odot n}$ (with n even and $\mathbf{a} \in V$) are symmetric positive semidefinite. This follows easily from the fact that $\mathbf{a}^{\odot n} \cdot \mathbf{v}^{\odot n} = (\mathbf{a} \cdot \mathbf{v})^n$, which is obviously non-negative for all $\mathbf{v} \in V$ if n is even. It follows that any weighted sum of such tensor powers, with positive weights, is symmetric positive definite. The cone generated by all tensor powers is thus a subset of the ordering cone of the Loewner order. It is not immediately clear, in the general case, whether it is a strict subset or whether the two cones are equal. However, for degree-two tensors it can be seen that the two cones are in fact the same (as in Fig. 1): it is well-known that any symmetric tensor of degree two can be written as a weighted sum of tensor squares of vectors that form an orthonormal basis (the eigendecomposition), and that such a tensor is positive semidefinite if and only if
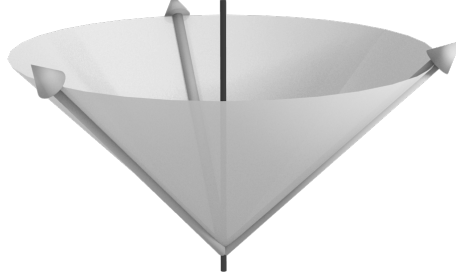
9

Figure 1: The cone spanned by tensor powers of degree two, where the underlying vector space is assumed to be a two-dimensional (Euclidean) space, visualized in a 3D representation of the space $(\mathbb{R}^2)^{\odot 2}$ (see [1] for the mapping). The vertical axis corresponds to the axis spanned by the degree-two identity tensor, the three arrows correspond to tensor squares of three equispaced vectors. Any tensor inside the cone would be considered larger than (or equal to) the tensor at the tip of the cone.

all weights are non-negative; consequently, any symmetric positive semidefinite tensor of degree two can be written as a non-negative combination of tensor squares, and is part of the cone spanned by all tensor squares.

We further characterize the cone through Lemma 1, which states that only scalar multiples of the identity tensor remain fixed under rotations (on V). Given that the identity tensor itself is positive semidefinite, and that the cone of positive semidefinite tensors is rotation invariant, we can thus consider the subspace spanned by the identity tensor to be the "centre" of the Loewner ordering cone.

Lemma 1. *The subspace of $V^{\odot n}$ spanned by the identity tensor $\mathbf{I}_n$ is precisely the set of tensors in $V^{\odot n}$ that are unaffected by rotations of the form $R^{\odot n}$, with R a rotation on V ($R \in SO(V)$ for short).*

Proof. To see that this is true, suppose we have some tensor $\mathbf{A}$ that remains unchanged under the aforementioned rotations. Without loss of generality we can assume that it must be some weighted sum of tensor powers of vectors $\mathbf{A} = \sum_{r \in \mathcal{R}} \lambda_r \, \mathbf{a}_r^{\odot n}$. Clearly, since it remains unchanged under rotations, we can rotate the individual tensor powers, sum them (with the appropriate weights) and get the original tensor back: $\mathbf{A} = \sum_{r \in \mathcal{R}} \lambda_r \, (R^{\odot n} \mathbf{a}_r^{\odot n})$. If we integrate over

all rotations — using a rotation invariant measure that integrates to one — then we get for some $c \in \mathbb{R}$

$$\mathbf{A} = \int_{R \in SO(V)} \sum_{r \in \mathcal{R}} \lambda_r \left(R^{\odot n} \mathbf{a}_r^{\odot n} \right) \mathrm{d}R$$

$$= \sum_{r \in \mathcal{R}} \lambda_r \int_{R \in SO(V)} R^{\odot n} \mathbf{a}_r^{\odot n} \, \mathrm{d}R = c \, \mathbf{I}_n.$$

The last, critical, identity follows from the fact that a rotation is an orthogonal transformation, and the assumed rotation invariance of the measure on $SO(V)$.

$\square$

Having defined the Loewner order for tensors, it can be used to define operations on tensor fields similar to dilation and erosion. This has been done by Burgeth et al. [1] for matrices and will be generalized below. However, as the minimum and maximum based on the Loewner order defined by Burgeth et al. do not give rise to a lattice these operations fail to satisfy most of the properties of dilations and erosions. Still, qualitatively there is a similarity to morphological operations, and in Section 4 we describe how these operations can formally be seen as a kind of approximation to proper morphological operators.

3.3. PDEs and eigendecomposition

In the current context the eigendecomposition of a symmetric matrix is most conveniently seen as a decomposition of a symmetric degree-two tensor into a (minimal) sum of rank-1 tensors: $\mathbf{A} = \sum_{k \in \mathcal{K}} \lambda_k \, \mathbf{a}_k^{\odot 2}$. For symmetric degree-two tensors (matrices) it is well-known that this decomposition always exists, is unique, and that the vectors $\{\mathbf{a}_k\}_{k \in \mathcal{K}}$ are orthogonal. For higher degree tensors the situation is a bit murkier [16, 22, 23], but we can still consider decompositions of the form $\mathbf{A} = \sum_{r \in \mathcal{R}} \lambda_r \, \mathbf{a}_r^{\odot n}$.

Burgeth et al. [1] consider a PDE-based scheme for constructing pseudo-dilations and -erosions (operators that are qualitatively similar to true dilations/erosions) on matrix fields, based on the eigendecomposition of symmetric matrices. The PDE $\frac{\partial p}{\partial t} = \|\nabla p\|$ that can be used for dilating a scalar field is applied to matrix fields by first rewriting the gradient norm as $\sqrt{dp(\mathbf{e}_1)^2 + \cdots}$,

11

$$\mathbb{R}^S \times \mathbb{R}^S \xdashrightarrow{\vee} \mathbb{R}^S$$
$$F \uparrow \qquad \qquad \downarrow P_+$$
$$V^{\odot n} \times V^{\odot n} \xrightarrow{\curlyvee_L} V^{\odot n}$$

$$\mathbb{R}^S \times \mathbb{R}^S \xrightarrow{\vee} \mathbb{R}^S$$
$$F \uparrow \qquad \qquad \downarrow F^+$$
$$V^{\odot n} \times V^{\odot n} \xdashrightarrow{\curlyvee_F} V^{\odot n}$$

Figure 2: Burgeth et al. [1] define a pseudo-join $\curlyvee_L$ directly on tensors, based on the Loewner order. We show that this pseudo-join can be interpreted as first lifting the tensors to their frame representation, then applying the regular join in that representation, and finally projecting back using P_+. That is, the diagram on the left commutes. Similarly, one could define a pseudo-join $\curlyvee_F$ based on least-squares backprojection (making the diagram on the right commute).

and then using a convention to extend the square and square root to symmetric matrices: a function h on the reals is extended to symmetric matrices by applying it to the eigenvalues of the eigendecomposition (part of an operator-algebraic view of symmetric matrices [13, 14]). For degree-two tensors we have $h(\mathbf{A}) = \sum_{k \in \mathcal{K}} h(\lambda_k) \mathbf{a}_k^{\otimes 2}$, where $\{\mathbf{a}_k\}_{k \in \mathcal{K}}$ is the set of eigenvectors of $\mathbf{A}$. We show below how this relates to performing the same kind of operations on a frame-based representation and demonstrate how this scheme can (to some degree) be generalized to higher degree tensors.

4. Frames and the Loewner order

Our frame-based method has very close ties with the Loewner order, but rather than leading to *pseudo*-dilations and -erosions, all the usual properties of morphological operators are preserved while working on the frame-based representation. Recall that a degree-n tensor $\mathbf{A}$ is considered positive semidefinite if (and only if) $\mathbf{A} \cdot \mathbf{v}^{\otimes n} \geq 0$ for all non-zero $\mathbf{v} \in V$. As we can obviously restrict ourselves to unit vectors in testing for positive semidefiniteness, a tensor is thus positive semidefinite if and only if all its coefficients in the frame are positive:

$$\mathbf{A} \leq_L \mathbf{B} \iff 0 \leq F_n(\mathbf{B} - \mathbf{A}) \iff F_n \mathbf{A} \leq F_n \mathbf{B}.$$
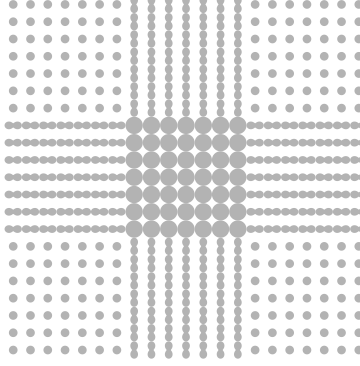
12

Figure 3: The (2D) tensor field used for filtering examples. Each glyph is effectively a polar plot of the frame coefficients of the tensors. When comparing the results in this paper to those obtained by Burgeth et al. [1], it is important to note that our experiment was carried out on a 2D dataset, not a 3D one (although the results are still quite similar).

The pseudo-join developed by Burgeth et al. [1], denoted by $\curlyvee_L$, finds the matrix with smallest trace that is an upper bound to both matrices (according to the Loewner order). To generalize this to tensors of arbitrary degree, we replace the trace by the inner product with the identity tensor. This is clearly equivalent for degree-2 matrices, and in general it gives the component parallel to the centre axis of the ordering cone (Lemma 1). We thus get:

$$\mathbf{A} \curlyvee_L \mathbf{B} = \underset{\mathbf{C} \in V^{\odot n}}{\arg\min} \, \mathbf{C} \cdot \mathbf{I}_n \text{ subject to } F_n\mathbf{A} \leq F_n\mathbf{C} \text{ and } F_n\mathbf{B} \leq F_n\mathbf{C}$$

$$= \underset{\mathbf{C} \in V^{\odot n}}{\arg\min} \, \mathbf{C} \cdot \mathbf{I}_n \text{ subject to } F_n\mathbf{A} \vee F_n\mathbf{B} \leq F_n\mathbf{C}.$$

To be able to view this result as a projection of some vector of frame coefficients $\mathbf{u} \in \mathbb{R}^S$ to the original tensor space $V^{\odot n}$ we take

$$P_+(\mathbf{u}) = \underset{\mathbf{A} \in V^{\odot n}}{\arg\min} \, \mathbf{A} \cdot \mathbf{I}_n \text{ subject to } \mathbf{u} \leq F_n\mathbf{A}.$$

Burgeth et al. implicitly show, through a reduction to the smallest enclosing circle problem [24], that this minimization is well-defined if $\mathbf{u}$ is representable by a join of lifted tensors. We will at least assume that $\mathbf{u}$ is bounded, as otherwise there is no feasible point. However, even then the minimization may not have a *unique* solution. Since the feasible set is clearly closed and convex, and the cost
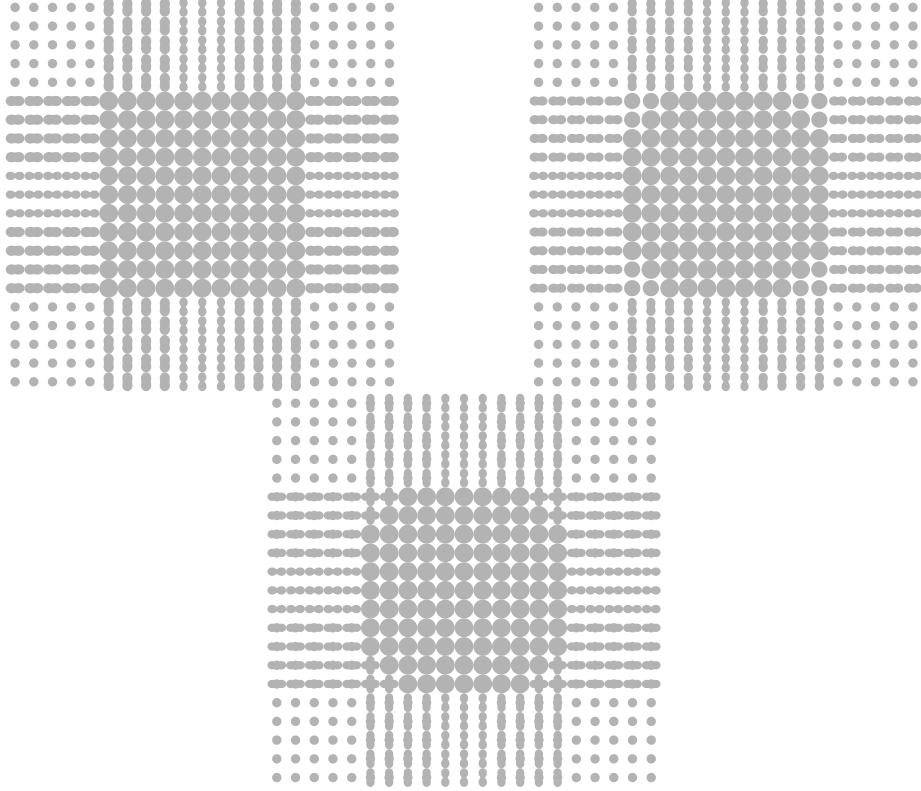
Figure 4: A dilation (with a ball of radius 2) is applied to the (2D) tensor field shown in Fig. 3, using (from left to right) the Loewner order, the frame representation (projecting back using F^+) and the frame representation without projecting back. Clearly, the left and middle columns are quite similar, with the differing projections only giving rise to small changes. The right-most column is again similar to the other two, but it has a higher angular resolution and can thus show crossings. Differences are most visible near the corners of the middle square.

function linear, we *can* conclude that there is always a (non-empty) convex set of optimal solutions. To remedy the situation one can thus use the following, slightly altered, definition:

$$T_+(\mathbf{u}) = \min_{\mathbf{A} \in V^{\odot n}} \mathbf{A} \cdot \mathbf{I}_n \text{ subject to } \mathbf{u} \leq F_n\mathbf{A},$$

$$P_+(\mathbf{u}) = \arg\min_{\mathbf{A} \in V^{\odot n}} \|F\mathbf{A} - \mathbf{u}\| \text{ subject to } \mathbf{A} \cdot \mathbf{I}_n \leq T_+(\mathbf{u}) \text{ and } \mathbf{u} \leq F_n\mathbf{A}.$$

As the feasible set is now the set of optimal solutions for the earlier formulation, it is still closed and convex. Also, the cost function is now *strictly* convex, so the modified P_+ has a unique solution, and it will be one of the optimal solutions for the original formulation. We are now in a position to redefine and generalize the pseudo-join introduced by Burgeth et al. using frames:

$$\mathbf{A} \curlyvee_L \mathbf{B} = \arg\min_{\mathbf{C} \in V^{\odot n}} \mathbf{C} \cdot \mathbf{I}_n \text{ subject to } F\mathbf{A} \vee F\mathbf{B} \leq F\mathbf{C}$$

$$= P_+(F\mathbf{A} \vee F\mathbf{B}).$$

Obviously, we can define an analogous projection P_- that gives the tensor with the largest "trace" that is a lower bound. We then observe that $\mathbf{A} \curlywedge_L \mathbf{B} = P_-(F\mathbf{A} \wedge F\mathbf{B})$, with '$\curlywedge_L$' being the pseudo-meet based on the Loewner order, as introduced by Burgeth et al. We can also define a rotation-invariant pseudo-join '$\curlyvee_F$' using the pseudo-inverse F^+ of F:

$$\mathbf{A} \curlyvee_F \mathbf{B} = F^+(F\mathbf{A} \vee F\mathbf{B}).$$

These results are summarized in Fig. 2. Figure 4 shows the effect of using different (pseudo-)joins on a dilation[3].

One of the advantages of our new take on filtering tensor data is that it shows that it can be advantageous to do as much as possible on the frame representation before projecting back. For example, if we were to construct a pseudo-opening using Burgeth et al.'s original pseudo-dilation and -erosion, then we effectively shuttle back and forth between the basis and frame representations twice. Computing the opening on the frame representation without

---

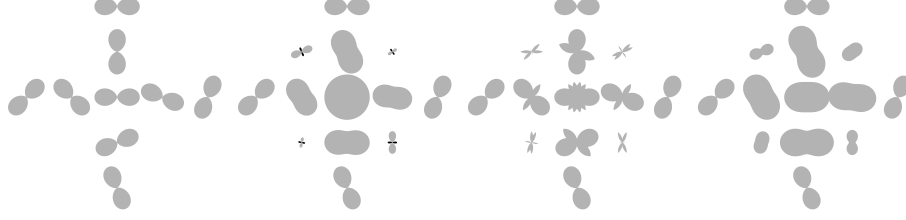[3]Code for Figs. 3 to 6 and 9 is available at `http://bit.ly/15VBe8w`.

15

Figure 5: From left to right: input, Burgeth et al.'s pseudo-closing, a closing in the frame representation, the Loewner-order-based projection (P_+) of the closing in the frame representation. In all cases, the structuring element is a 3-by-3 cross. Notice that the pseudo-closing results in an isotropic tensor in the center of the image, and that it does not preserve the positive semidefiniteness of the matrices (indicated by the black parts on the glyphs to the north west (NW), NE, SE and SW of the center glyph).
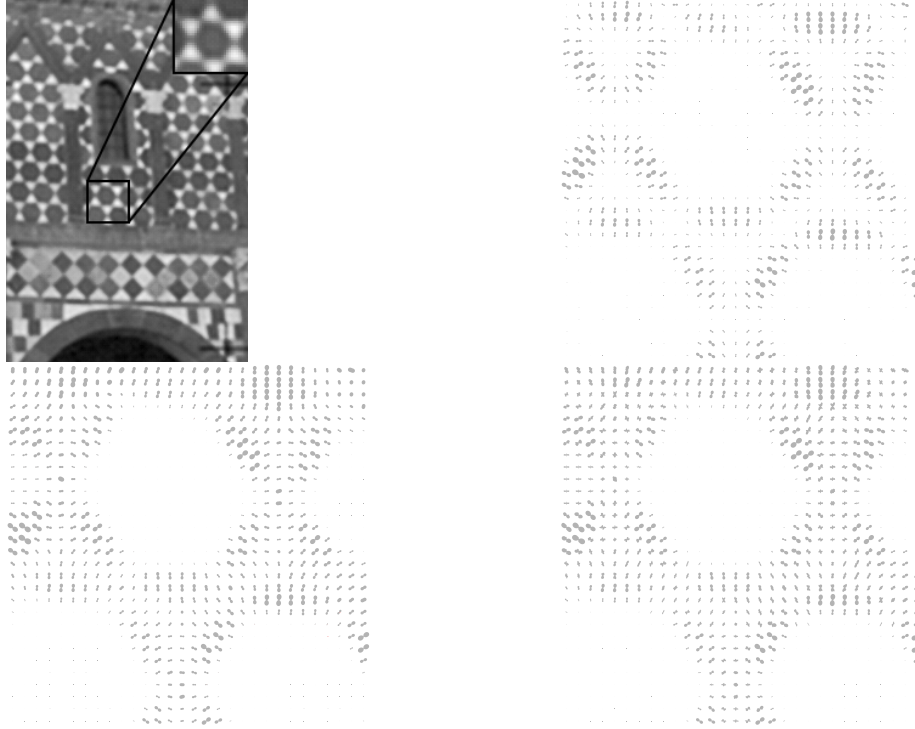

Figure 6: From left to right: the original image with magnified inset of the part used, the input structure tensor field (see [4, 25]), the structure tensor field dilated using the Loewner order, and the structure tensor field dilated using frames. The structuring element was a ball of radius 3.

$_{321}$ going back to the basis representation and then lifting ensures that what we
$_{322}$ compute is in fact an opening until we project back to the basis representation.
$_{323}$ Figure 5 and Fig. 6 compare both approaches. In a sense we acknowledge that
$_{324}$ certain operations simply do not make sense on the original representation, so
$_{325}$ we construct a new representation on which they do make sense, and only when
$_{326}$ absolutely necessary do we go back (minimizing the error we make in doing so).
$_{327}$ On the other hand, using P_+ to project back to the original tensor space,
$_{328}$ as implicitly done in the work of Burgeth et al. [1], may also have its merits,
$_{329}$ especially if used only as the last step in processing a tensor field. In particular,
$_{330}$ it guarantees that the result is in fact an upper bound of the frame-based result.
$_{331}$ This might be important for fibre tracking for example, as that typically relies
$_{332}$ mostly on the direction and magnitude of the highest response of a tensor.

$_{333}$ **5. PDE-based morphology**

$_{334}$ Apart from the Loewner-order-based approach, Burgeth et al. also give an
$_{335}$ approach based on PDEs. It was made apparent that there is a link between
$_{336}$ these two approaches, but this link was not explored much. Here we show that
$_{337}$ the two approaches are indeed very deeply related to each other, as well as to
$_{338}$ our frame-based method.

$_{339}$ The traditional PDE for a dilation with a disk is based on the gradient
$_{340}$ magnitude: $\frac{\partial p}{\partial t} = \|\nabla p\|$, with p a continuously differentiable scalar field on
$_{341}$ some domain Ω ($p : \Omega \to \mathbb{R}$). It is well-known that the gradient magnitude
$_{342}$ corresponds to the value of the largest directional derivative. We denote the
$_{343}$ directional derivative (field) in direction $\mathbf{s}$ using the differential as $\mathrm{d}p(\mathbf{s})$, and
$_{344}$ we thus get

$_{345}$
$$\frac{\partial p}{\partial t} = \|\nabla p\| = \bigvee_{\mathbf{s} \in S_V} \mathrm{d}p(\mathbf{s}).$$

$_{346}$ By analogy, supposing we have lifted a *tensor* field $f : \Omega \to V^{\odot n}$ to the frame
$_{347}$ representation $g = F_n f$ (where F_n is applied to all positions separately), for

17

$$\begin{array}{ccccccc}
\mathcal{M} & \xrightarrow{\mathrm{d}\cdot(\mathbf{e}_k)} & \mathcal{M}^{\mathcal{K}} & \xrightarrow{\;.2\;} & \mathcal{M}^{\mathcal{K}} & & \\
\Big\uparrow F & & F\Big\uparrow P_+ & & \Big\uparrow P_+ & & \\
(f)\;\; \mathcal{L} & \xrightarrow{\mathrm{d}\cdot(\mathbf{e}_k)} & \mathcal{L}^{\mathcal{K}} & \xrightarrow{\;.2\;} & \mathcal{L}^{\mathcal{K}} & \xrightarrow{\Sigma}\;\mathcal{L}\;\xrightarrow{\sqrt{\cdot}}\;\mathcal{L}\;(\tfrac{\partial f}{\partial t}) \\
& & & & \Big\downarrow F & F\Big\downarrow P_- \quad \Big\downarrow P_- \\
& & & & \mathcal{M}^{\mathcal{K}} & \xrightarrow{\Sigma} \mathcal{M} \xrightarrow{\sqrt{\cdot}} \mathcal{M}
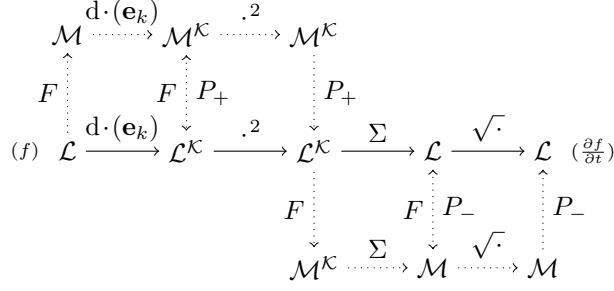\end{array}$$

Figure 7: Overview of Burgeth et al.'s method for computing a tensorial equivalent of $\frac{\partial p}{\partial t} = \|\nabla p\|$. This diagram commutes only in the sense that all paths starting at f and ending at $\frac{\partial f}{\partial t}$ give the same result (more specifically, $F \circ P_\pm$ is not equivalent to the identity operator, and the square (root) does not commute with the projection operator). Here $\mathcal{L}$ is the set of functions from Ω to $V^{\odot n}$ and $\mathcal{M}$ is the set of functions from Ω to $\mathbb{R}^S$ (operations are taken point-wise).

$g : \Omega \to \mathbb{R}^S$ we would have

$$\frac{\partial g}{\partial t} = \bigvee_{\mathbf{s} \in S_V} \mathrm{d}g(\mathbf{s}),$$

$$\frac{\partial g_{\mathbf{a}}}{\partial t} = \left[\bigvee_{\mathbf{s} \in S_V} \mathrm{d}g(\mathbf{s}) \right]_{\mathbf{a}} = \bigvee_{\mathbf{s} \in S_V} \mathrm{d}g(\mathbf{s})_{\mathbf{a}} = \bigvee_{\mathbf{s} \in S_V} \mathrm{d}g_{\mathbf{a}}(\mathbf{s}).$$

Since $g_{\mathbf{a}}$ itself is a scalar field, we have $\frac{\partial g_{\mathbf{a}}}{\partial t} = \|\nabla g_{\mathbf{a}}\| = \sqrt{\mathrm{d}g_{\mathbf{a}}(\mathbf{e_1})^2 + \mathrm{d}g_{\mathbf{a}}(\mathbf{e_2})^2 + \cdots}$, and by extension

$$\frac{\partial g}{\partial t} = \sqrt{\mathrm{d}g(\mathbf{e_1})^2 + \mathrm{d}g(\mathbf{e_2})^2 + \cdots}.$$

Interestingly, Burgeth et al. [1] directly lift $\|\nabla p\|$ to matrix fields by writing the norm of the gradient as the square root of the squared partial derivatives and applying this formula to matrices using their convention for lifting functions on scalars to functions on matrices (by applying the function to the eigenvalues of the matrix). Using this convention we thus get:

$$\frac{\partial f}{\partial t} = \sqrt{\mathrm{d}f(\mathbf{e_1})^2 + \mathrm{d}f(\mathbf{e_2})^2 + \cdots}.$$

So how does this compare to the result we got for g? Using Theorem 1 below we can see that, similar to how the Loewner-order-based join $\mathbf{A} \curlyvee_L \mathbf{B}$ is equal to
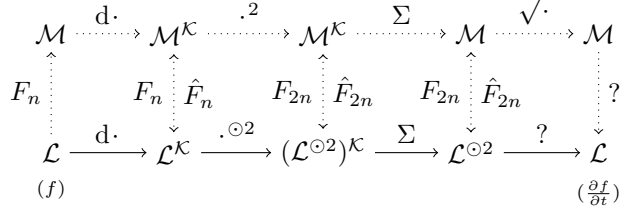
18

$$\mathcal{M} \xdashrightarrow{\;d\cdot\;} \mathcal{M}^{\mathcal{K}} \xdashrightarrow{\;\cdot^2\;} \mathcal{M}^{\mathcal{K}} \xdashrightarrow{\;\Sigma\;} \mathcal{M} \xdashrightarrow{\;\sqrt{\cdot}\;} \mathcal{M}$$

$$F_n \quad\quad F_n \;\; \hat{F}_n \quad\quad F_{2n} \;\; \hat{F}_{2n} \quad\quad F_{2n} \;\; \hat{F}_{2n} \quad\quad ?$$

$$\mathcal{L} \xrightarrow{\;d\cdot\;} \mathcal{L}^{\mathcal{K}} \xrightarrow{\;\cdot^{\odot 2}\;} (\mathcal{L}^{\odot 2})^{\mathcal{K}} \xrightarrow{\;\Sigma\;} \mathcal{L}^{\odot 2} \xrightarrow{\;?\;} \mathcal{L}$$

$$(f) \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \left(\tfrac{\partial f}{\partial t}\right)$$

Figure 8: Using a tensor square instead of squaring the eigenvalues makes for a much closer correspondence between the frame-based computation and the computation performed directly on the tensors. In fact, only the square root clearly cannot be represented exactly using finite-degree tensors. Note that the same kind of commutation rule applies as in Fig. 7. Here $\hat{F}$ stands for an arbitrary left-inverse of F, so $\hat{F} \circ F = \text{id}$.

$P_+(F\mathbf{A} \vee F\mathbf{B})$, the square of a matrix $\mathbf{A}^2$ equals $P_+([F\mathbf{A}]^2)$, while the square root of a (positive-semidefinite) matrix $\sqrt{\mathbf{A}}$ equals $P_-(\sqrt{F\mathbf{A}})$. Figure 7 shows the correspondence between the tensor- and frame-based computations in detail (also using F^+ instead of $P_\pm$). It shows that Burgeth et al.'s computation essentially goes back and forth between the frame representation twice, which is obviously not ideal in terms of approximating the frame-based result. Figure 9 shows examples of filtering Fig. 3 using PDEs based on both Burgeth et al.'s PDEs on matrices, as well as our frame-based methods.

To get a more direct correspondence with what was derived for the frame representation, one can replace the square in the original definition of Burgeth et al. by a tensor square (see Fig. 8). However, it is then not immediately obvious how to compute the square root on the tensor-based representation. To have a close analogy with the Loewner-order-based pseudo-dilation, one might desire a square-root-like operation on a degree-$2n$ tensor $\mathbf{A}$ that gives a degree-n tensor $\mathbf{B}$ with the smallest inner product with the identity tensor, such that $F_n\mathbf{B} \leq \sqrt{F_{2n}\mathbf{A}}$. Alternatively, a least-squares solution based on $F_n^+\sqrt{F_{2n}\mathbf{A}}$ might be attractive.

To work up to showing that $h(\mathbf{A})$ equals $P_+(h(F\mathbf{A}))$ for any symmetric degree-two tensor $\mathbf{A}$ and convex h, we first establish in Lemma 2 and Corollary 1 that the frame coefficients of a degree-two tensor are always expressible as convex
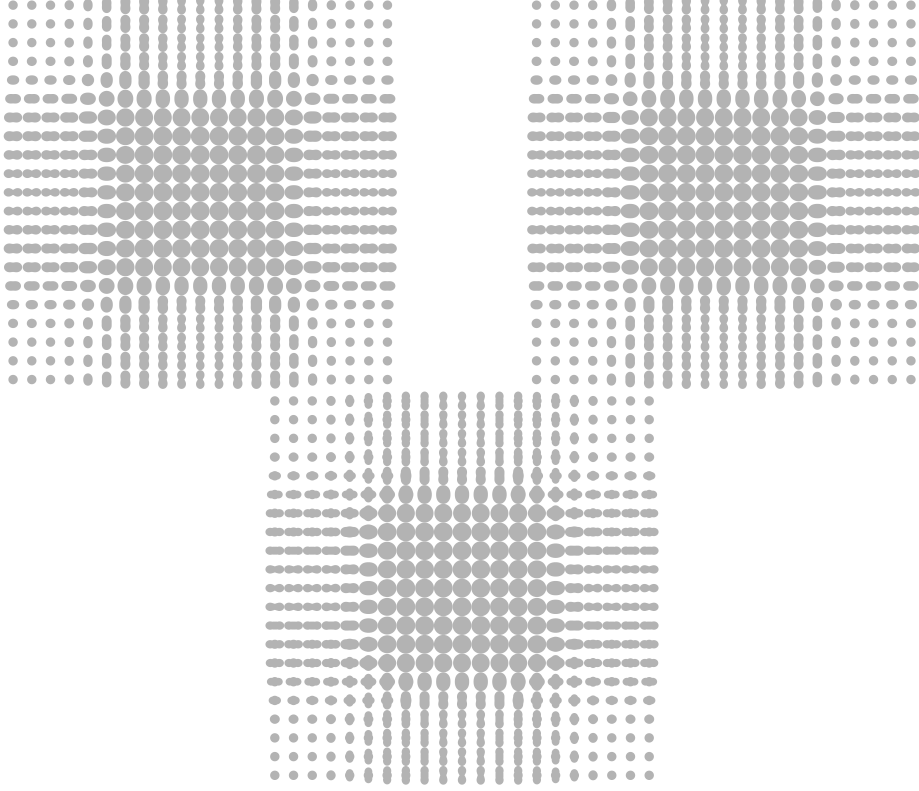
19

Figure 9: The same as Fig. 4, except now based on a (very simple) PDE solver, causing some blurring.

combinations of its eigenvalues. Then, in Lemma 3 and Theorem 1, we show that this indeed allows us to conclude that $h(\mathbf{A})$ equals $P_+(h(F\mathbf{A}))$ for $\mathbf{A} \in V^{\odot 2}$ and convex h.

Lemma 2. *Suppose* $\mathbf{A}$ *is a symmetric tensor of even degree* n, *with a decomposition* $\mathbf{A} = \sum_{r \in \mathcal{R}} \lambda_r \, \mathbf{a}_r^{\odot n}$ *such that[4]* $\sum_{r \in \mathcal{R}} \mathbf{a}_r^{\odot n} = \mathbf{I}_n$. *We then have for any unit vector* $\mathbf{s} \in S_V$ *that* $\mathbf{s}^{\odot n} \cdot \mathbf{A}$ *is a convex combination of the values* $\{\lambda_r\}_{r \in \mathcal{R}}$.

---

[4]The $\mathbf{a}_r$ need not be mutually orthogonal, nor of unit length.

Proof. Clearly, we have

$$\mathbf{s}^{\odot n} \cdot \mathbf{A} = \mathbf{s}^{\odot n} \cdot \left(\sum_{r \in \mathcal{R}} \lambda_r \, \mathbf{a}_r^{\odot n} \right)$$

$$= \sum_{r \in \mathcal{R}} \lambda_r (\mathbf{s}^{\odot n} \cdot \mathbf{a}_r^{\odot n}) = \sum_{r \in \mathcal{R}} \lambda_r (\mathbf{s} \cdot \mathbf{a}_r)^n.$$

The question now is: are the coefficients $\{(\mathbf{s} \cdot \mathbf{a}_r)^n\}_{r \in \mathcal{R}}$ all non-negative, and do they add up to one? That the coefficients are non-negative is easily seen: as n is assumed to be even, $(\mathbf{s} \cdot \mathbf{a}_r)^n$ clearly must be non-negative. Since $\mathbf{s}$ is a unit vector and we assumed that $\sum_{r \in \mathcal{R}} \mathbf{a}_r^{\odot n} = \mathbf{I}_n$, we also have

$$\sum_{r \in \mathcal{R}} (\mathbf{s} \cdot \mathbf{a}_r)^n = \sum_{r \in \mathcal{R}} \mathbf{s}^{\odot n} \cdot \mathbf{a}_r^{\odot n}$$

$$= \mathbf{s}^{\odot n} \cdot \left(\sum_{r \in \mathcal{R}} \mathbf{a}_r^{\odot n} \right) = \mathbf{s}^{\otimes n} \cdot \mathbf{I}_n = \|\mathbf{s}\|^n = 1.$$

This concludes the proof that $\mathbf{s}^{\odot n} \cdot \mathbf{A}$ is a convex combination of the values $\{\lambda_r\}_{r \in \mathcal{R}}$ referred to in the lemma for any $\mathbf{s} \in S_V$. $\qquad\square$

Corollary 1. *Given a symmetric degree-two tensor* $\mathbf{A}$, $\mathbf{s}^{\odot 2} \cdot \mathbf{A}$ *is a convex combination of the eigenvalues of* $\mathbf{A}$ *for any unit vector* $\mathbf{s}$.

Proof. The eigendecomposition of a symmetric degree-two tensor (matrix) can always be considered to give an orthonormal set of eigenvectors, and such a set clearly sums to the degree-two identity tensor. The corollary now follows from Lemma 2. $\qquad\square$

Lemma 3. *Given a convex function* $h : \mathbb{R} \to \mathbb{R}$, *the lifted versions* $h : V^{\odot 2} \to V^{\odot 2}$ *and* $h : \mathbb{R}^S \to \mathbb{R}^S$, *and a symmetric degree-two tensor* $\mathbf{A}$, $Fh(\mathbf{A})$ *bounds* $h(F\mathbf{A})$ *from above and there is no tensor* $\mathbf{B}$ *with strictly smaller trace than* $h(\mathbf{A})$ *for which* $F\mathbf{B}$ *bounds* $h(F\mathbf{A})$ *from above. The same holds for a concave function (with the inequalities reversed).*

Proof. First we prove that $Fh(\mathbf{A})$ bounds $h(F\mathbf{A})$ from above, and then that there is no other tensor with strictly smaller trace than $h(\mathbf{A})$ that does the same.

21

$_{417}$ To this end, we first note that $(Fh(\mathbf{A}))_\mathbf{s}$ matches $h(F\mathbf{A})_\mathbf{s}$ for any eigenvector $\mathbf{s}$

$_{418}$ of $\mathbf{A}$. Also, for all other vectors, $Fh(\mathbf{A})$ is larger than or equal to $h(F\mathbf{A})$, due

$_{419}$ to the convexity of h and Corollary 1.

$_{420}$ It now remains for us to see whether there can be another tensor $\mathbf{B}$ with

$_{421}$ strictly smaller trace such that $F\mathbf{B}$ is an upper bound of $h(F\mathbf{A})$. It can be

$_{422}$ shown that this is not the case, by considering the (degree-two) identity tensor

$_{423}$ to be the sum of the tensor squares of the eigenvectors of $\mathbf{A}$, and recalling

$_{424}$ that the trace of a degree-two tensor can be defined as its inner product with

$_{425}$ the identity tensor. $F\mathbf{B}$ must be greater than or equal to $h(F\mathbf{A})$, and $h(F\mathbf{A})$

$_{426}$ matches $Fh(\mathbf{A})$ for the eigenvectors of $\mathbf{A}$. So clearly the trace of $\mathbf{B}$ must be at

$_{427}$ least the trace of $h(\mathbf{A})$. This concludes the proof.

$_{428}$ $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

$_{429}$ **Theorem 1.** *Given a convex continuously-differentiable function $h : \mathbb{R} \to \mathbb{R}$, the*

$_{430}$ *lifted versions $h : V^{\odot 2} \to V^{\odot 2}$ and $h : \mathbb{R}^S \to \mathbb{R}^S$, and a symmetric degree-two*

$_{431}$ *tensor $\mathbf{A}$, $P_+(h(F\mathbf{A}))$ is equal to $h(\mathbf{A})$.*

$_{432}$ *Proof.* $P_+(h(F\mathbf{A}))$ will have the same trace as $h(\mathbf{A})$. This follows from Lemma 3

$_{433}$ and the observation that the trace of a degree-two tensor can be defined as its

$_{434}$ inner product with the identity tensor. So what remains is to show that there is

$_{435}$ only one tensor with that trace whose frame representation is an upper bound

$_{436}$ for $h(F\mathbf{A})$.

$_{437}$ Suppose that there is some other tensor $\mathbf{B} = P_+(h(F\mathbf{A}))$ with the same

$_{438}$ trace as $h(\mathbf{A})$, such that $F\mathbf{B} \geq h(F\mathbf{A})$. As discussed above, $(F\mathbf{B})_\mathbf{s}$ must equal

$_{439}$ $h(F\mathbf{A})_\mathbf{s}$ for any eigenvector $\mathbf{s}$ of $\mathbf{A}$. Clearly, $(F\mathbf{B})_\mathbf{s} = \mathbf{B} \cdot \mathbf{s}^{\odot 2}$ must then also be

$_{440}$ tangent to $h(F\mathbf{A})_\mathbf{s} = h(\mathbf{A} \cdot \mathbf{s}^{\odot 2})$ for any eigenvector $\mathbf{s}$ to have $F\mathbf{B} \geq h(F\mathbf{A})$.

$_{441}$ Since h is continuously-differentiable, we have $\nabla_\mathbf{v} h(\mathbf{A} \cdot \mathbf{v}^{\odot 2}) = 2\, h'(\mathbf{A} \cdot \mathbf{v}^{\odot 2})\, \mathbf{A}\mathbf{v}$

$_{442}$ (with $\mathbf{v} \in V$). So we have $\nabla_\mathbf{v} h(\mathbf{A} \cdot \mathbf{v}^{\odot 2}) = c\, \mathbf{v}$ ($c \in \mathbb{R}$) if and only if $\mathbf{v}$ is

$_{443}$ an eigenvector of $\mathbf{A}$. Something similar holds for $\mathbf{B}$, and we thus see that $\mathbf{B}$,

$_{444}$ $\mathbf{A}$ and $h(\mathbf{A})$ all have the same eigenvectors. $\mathbf{B}$ must then also have the same

$_{445}$ eigenvalues as $h(\mathbf{A})$, and is thus identical to $h(\mathbf{A})$. This concludes the proof. $\square$

It is currently unclear whether or not (or at least how) the above generalizes to higher degree tensors. This mostly has to do with the fact that eigenvalue theory for higher degree tensors is not yet as well-developed as eigenvalue theory for degree-two tensors. The recent survey done by Chang et al. [26] is a nice starting point for the reader interested in this subject.

6. Approximation quality

The reader might wonder how good the approximations made by Burgeth et al.'s and our methods are. What kind of error guarantees can be given? Below we examine the "error" made by functions lifted to degree-two tensors.

Lemma 4. *Given a convex function* $h : \mathbb{R} \to \mathbb{R}$, *the lifted versions* $h : V^{\odot 2} \to V^{\odot 2}$ *and* $h : \mathbb{R}^S \to \mathbb{R}^S$, *and a degree-two tensor* $\mathbf{A}$ *with an eigendecomposition* $\mathbf{A} = \sum_{k \in \mathcal{K}} \lambda_k \, \mathbf{a}_k^{\odot 2}$ *such that* $\sum_{k \in \mathcal{K}} \mathbf{a}_k^{\odot 2} = \mathbf{I}_2$, *the error* $|Fh(\mathbf{A}) - h(F\mathbf{A})|$ *is bounded from above by*

$$\max_{0 \leq t \leq 1} (1 - t) \, h(\lambda_-) + t \, h(\lambda_+) - h((1 - t) \, \lambda_- + t \, \lambda_+),$$

with λ_- *the smallest* λ_k *and* λ_+ *the largest* λ_k. *For concave functions we only need to change the sign of the function being maximized.*

Proof. Note that $Fh(\mathbf{A}) - h(F\mathbf{A}) \geq 0$ because of the convexity of h and Lemma 3. For any $\mathbf{s} \in S_V$, $h(F\mathbf{A})_{\mathbf{s}}$ equals h applied to a convex combination of the eigenvalues of $\mathbf{A}$, while $[Fh(\mathbf{A})]_{\mathbf{s}}$ equals a convex combination (with the same weights) of the eigenvalues transformed by h (see Lemma 2). The question thus is: what is the largest possible difference between $\sum_{k \in \mathcal{K}} w_k \, h(\lambda_k)$ and $h(\sum_{k \in \mathcal{K}} w_k \, \lambda_k)$, where the w_k are non-negative and sum up to one?

Without loss of generality, take $\sum_{k \in \mathcal{K}} w_k \, \lambda_k$ to equal $(1 - t) \, \lambda_- + t \, \lambda_+$, then

$$\sum_{k \in \mathcal{K}} w_k \, h(\lambda_k) - h(\sum_{k \in \mathcal{K}} w_k \, \lambda_k) \leq$$

$$(1 - t) \, h(\lambda_-) + t \, h(\lambda_+) - h((1 - t) \, \lambda_- + t \, \lambda_+),$$

23

because $h(\sum_{k \in \mathcal{K}} w_k \, \lambda_k) = h((1 - t)\, \lambda_- + t\, \lambda_+)$ and

$$\sum_{k \in \mathcal{K}} w_k \, h(\lambda_k) \leq (1 - t)\, h(\lambda_-) + t\, h(\lambda_+).$$

The latter can be seen to be true by replacing $h(\lambda_k)$ by a convex combination $(1 - t_k)\, h(\lambda_-) + t_k \, h(\lambda_+)$ of $h(\lambda_-)$ and $h(\lambda_+)$ such that the same convex combination of λ_- and λ_+ equals λ_k. Because h is convex this can only create something greater than or equal to $\sum_{k \in \mathcal{K}} w_k \, h(\lambda_k)$. And we see that this sum $(\sum_{k \in \mathcal{K}} w_k \, [(1 - t_k)\, h(\lambda_-) + t_k \, h(\lambda_+)])$ must indeed equal $(1 - t)\, h(\lambda_-) + t\, h(\lambda_+)$, because $\sum_{k \in \mathcal{K}} w_k \, \lambda_k = \sum_{k \in \mathcal{K}} w_k \, [(1 - t_k)\, \lambda_- + t_k \, \lambda_+]$ equals $(1 - t)\, \lambda_- + t\, \lambda_+$. This means that we can limit ourselves to examining $(1 - t)\, h(\lambda_-) + t\, h(\lambda_+) - h((1 - t)\, \lambda_- + t\, \lambda_+)$ with $0 \leq t \leq 1$, rather than the more general $\sum_{k \in \mathcal{K}} w_k \, h(\lambda_k) - h(\sum_{k \in \mathcal{K}} w_k \, \lambda_k)$.

It now naturally follows that the error $F h(\mathbf{A}) - h(F \mathbf{A})$ is bounded from above by $\max_{0 \leq t \leq 1} (1 - t)\, h(\lambda_-) + t\, h(\lambda_+) - h((1 - t)\, \lambda_- + t\, \lambda_+)$, which concludes the proof. $\qquad \square$

Corollary 2. *The error $|F \mathbf{A}^2 - (F \mathbf{A})^2|$ is bounded from above by $1/4\, (\lambda_+ - \lambda_-)^2$.*

Proof. We simply take the expression for the error bound found in Lemma 4, fill in squaring for h and simplify:

$$(1 - t)\, \lambda_-^2 + t\, \lambda_+^2 - ((1 - t)\, \lambda_- + t\, \lambda_+)^2$$

$$= (1 - t)\, \lambda_-^2 + t\, \lambda_+^2 - (1 - t)^2\, \lambda_-^2$$

$$- 2\, (1 - t)\, t\, \lambda_- \, \lambda_+ - t^2\, \lambda_+^2$$

$$= (1 - t - (1 - t)^2)\, \lambda_-^2 + (t - t^2)\, \lambda_+^2 - 2\, (t - t^2)\, \lambda_- \, \lambda_+$$

$$= (t - t^2)\, \lambda_-^2 + (t - t^2)\, \lambda_+^2 - 2\, (t - t^2)\, \lambda_- \, \lambda_+$$

$$= (t - t^2)\, (\lambda_-^2 - 2\, \lambda_- \, \lambda_+ + \lambda_+^2)$$

$$= (t - t^2)\, (\lambda_+ - \lambda_-)^2.$$

Differentiating this with respect to t gives a zero at $t = 1/2$, with a value of $1/4\, (\lambda_+ - \lambda_-)^2$. $\qquad \square$

24

Corollary 3. *The error* $|\sqrt{F\mathbf{A}} - F\sqrt{\mathbf{A}}|$ *is bounded from above by*[5] $\frac{(\sqrt{\lambda_+} - \sqrt{\lambda_-})^2}{4\,(\sqrt{\lambda_-} + \sqrt{\lambda_+})}$ *and* $\frac{\sqrt{\lambda_+}}{4}$.

Proof. We take the expression for the error bound found in Lemma 4, negate it because the square root function is concave, fill in the square root for h, and take $\lambda_- = a^2$ and $\lambda_+ = b^2$ with $0 \le a < b$ (we can assume without loss of generality that a and b are non-negative, and if they are both zero the statement would be trivially true):

$$\sqrt{(1-t)\,\lambda_- + t\,\lambda_+} - \left((1-t)\,\sqrt{\lambda_-} + t\,\sqrt{\lambda_+}\right)$$
$$= \sqrt{(1-t)\,a^2 + t\,b^2} - \left((1-t)\,a + t\,b\right).$$

Differentiating this with respect to t, equating to zero and solving for t gives:

$$\frac{\partial}{\partial t}\left[\sqrt{(1-t)\,a^2 + t\,b^2}\right.$$
$$\left. - \left((1-t)\,a + t\,b\right)\right] = 0$$
$$\frac{(b-a)(b+a)}{2\sqrt{(1-t)\,a^2 + t\,b^2}} = (b-a)$$
$$\frac{1}{2}(b+a) = \sqrt{(1-t)\,a^2 + t\,b^2}$$
$$\frac{1}{4}(b+a)^2 = (1-t)\,a^2 + t\,b^2$$
$$\frac{1}{4}(b^2 + 2\,a\,b + a^2) = a^2 - t\,(a^2 - b^2)$$
$$t\,(a-b)\,(a+b) = \frac{1}{4}(3\,a + b)\,(a-b)$$
$$t = \frac{3\,a + b}{4\,(a+b)}.$$

---

[5]For the square root to make sense (in the current context), $\mathbf{A}$ should be positive semidefinite.

Filling back in gives:

$$\sqrt{(1 - \frac{3\,a+b}{4\,(a+b)})\,a^2 + \frac{3\,a+b}{4\,(a+b)}\,b^2}$$

$$- \left[(1 - \frac{3\,a+b}{4\,(a+b)})\,a + \frac{3\,a+b}{4\,(a+b)}\,b\right]$$

$$= \sqrt{a^2 + \frac{(3\,a+b)\,(b-a)\,(b+a)}{4\,(a+b)}}$$

$$- \left[a + \frac{(3\,a+b)(b-a)}{4\,(a+b)}\right]$$

$$= \sqrt{\frac{1}{4}(b+a)^2} - \left[a + \frac{(3\,a+b)(b-a)}{4\,(a+b)}\right]$$

$$= \frac{1}{2}(b+a) - \left[a + \frac{(3\,a+b)(b-a)}{4\,(a+b)}\right]$$

$$= \frac{1}{2}(b-a) - \frac{(3\,a+b)(b-a)}{4\,(a+b)}$$

$$= \frac{(b-a)^2}{4\,(a+b)}.$$

It is not too difficult to see that $\frac{(b-a)^2}{4\,(a+b)}$ is bounded from above by $\frac{b^2}{4\,b} = \frac{b}{4}$ (given that $0 \leq a < b$). Finally, filling in $\sqrt{\lambda_-}$ for a and $\sqrt{\lambda_+}$ for b, we get the first error bound ($\frac{(\sqrt{\lambda_+} - \sqrt{\lambda_-})^2}{4\,(\sqrt{\lambda_-} + \sqrt{\lambda_+})}$), as well as the looser but simpler bound $\frac{\sqrt{\lambda_+}}{4}$. $\quad\square$

Note that Lemma 4 and its corollaries would essentially apply to higher degree tensors as well, if we were to generalize the construction for lifting h using Lemma 2. However, since for higher degree tensors the λ's might not correspond directly to any frame coefficients, and we currently do not know how large the gap would be, it seems to make little sense to generalize Lemma 4 at this time.

For the moment, we have not derived bounds similar to those above for least-squares projection back to the original space. However, it is to be expected that the bounds would be the same or better.

In addition to the error bounds given above, it is also interesting to note that the Loewner-order-based projections may give qualitatively different answers. For example, in Fig. 10 we see that P_+ and F^+ give orthogonal primary
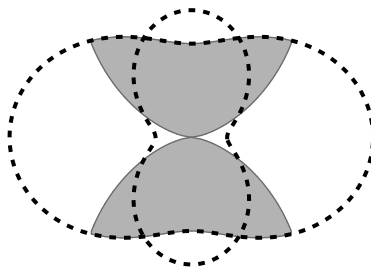
26

Figure 10: Projections using P_+ and F^+ of the meet (in the frame representation) of two tensors. The meet is shown as a solid grey glyph, the projections are shown using dashed lines. The horizontally oriented projection was done using P_+, the vertically oriented projection using F^+. Note that normally P_- would be used for projecting a meet back to the original tensor space, but this kind of situation can occur when doing an opening or closing on the frame representation and projecting back afterwards.

directions (the eigenvectors associated with the largest eigenvalues in either case are orthogonal).

7. Discussion

To our knowledge, the methods developed by Burgeth et al. [1] are the only approaches to tensor morphology (apart from ours) that attempt to directly generalize lattice-based mathematical morphology to tensors and tensor fields. Reinterpreting both these approaches as specific applications of our frame-based method allows us to view all existing approaches to tensor morphology — at least those based on directly building a lattice(-like) structure on tensor spaces — within a single framework.

So when is what approach called for? Conceptually, we typically prefer our frame-based approach, as the coefficients operated upon are often meaningful, and it allows for the most direct generalization of traditional morphological operators. It can also be viewed as a generalization of the work done by Burgeth et al. [1], and it even fits in very well with the work done by Duits et al. [12] (effectively the same representation is used). In practice, it is also often a lot easier to work with a frame-based approach: some matrix multiplications are

enough to go to and from the frame representation (assuming you work with a finite subset of the frame vectors), and the frame representation itself can often be processed using existing greyscale operators.

When it is important that the result of a dilation (for example) is actually an upper bound of the original signal, then using P_+ either explicitly or implicitly is of course preferable over using least-squares projection. It might be interesting to develop other projections with similar guarantees though.

PDE-based morphology in general has as an advantage that it allows for subpixel accuracy [27], while typically introducing at least some blur. Also, PDE-based methods allow for arbitrary radii when using a disk-shaped structuring element, while other methods always have to approximate a disk by some discrete set of pixels. The current work clears up how the PDEs developed by Burgeth et al. can be interpreted to compute an approximation to a lattice-theoretic dilation/erosion, and how this approximation might be improved.

The current work also exposes an interesting connection between recent work done by Burgeth and Kleefeld [28] on colour morphology and our (concurrent) work done on colour morphology [3]. Apart from some differences in the colour spaces used and similar technicalities, the Loewner-order-based approach for colour morphology developed by Burgeth and Kleefeld is very similar to the hue invariant frame approach we developed. In fact, if one would identify the "grey axis" with the centre of the Loewner order, the main difference would be that the Loewner ordering cone has a slightly different angle with respect to its centre. In future work it might make sense to take a closer look at this connection, and perhaps evaluate using F^+ rather than $P_{\pm}$ to go back to the original colour space.

8. Conclusion

We have shown that two earlier methods for tensor morphology developed by Burgeth et al. [1] can be interpreted within our frame-based approach [2–4]. In particular, both methods are equivalent to using our method with a particular

28

projection back to the original space. Also, the PDE-based methods implicitly go back and forth between the frame-based representation several times.

We have also given a quantitative analysis of the "error" made in Burgeth et al. [1]'s approaches, compared to sticking to a frame-based representation. It remains to be seen how a least-squares projection compares exactly, but by construction its (least-squares) error should obviously be smaller. It is interesting to note that the relative error remains bounded, albeit not particularly small (up to about a quarter).

Conceptually, the frame-based approach has the advantage of fitting better within the traditional morphological theory. On the other hand, the approaches based on the Loewner order can work directly on the original tensors and provide lower/upper bounds in the original tensor space. In a sense the current paper shows how we might have the best of both worlds by showing how the Loewner-order-based approaches can be seen as instances of the frame-based approach. This has already led to the identification of some potential improvements on both sides.

[1] B. Burgeth, A. Bruhn, S. Didas, J. Weickert, M. Welk, Morphology for matrix data: Ordering versus PDE-based approach, Image Vis. Comput. 25 (4) (2007) 496–511, ISSN 02628856, doi:10.1016/j.imavis.2006.06.002.

[2] J. J. van de Gronde, J. B. T. M. Roerdink, Group-invariant colour morphology based on frames, IEEE Trans. Image Process. Accepted.

[3] J. J. van de Gronde, J. B. T. M. Roerdink, Group-Invariant Frames for Colour Morphology, in: C. L. Luengo Hendriks, G. Borgefors, R. Strand (Eds.), Mathematical Morphology and Its Applications to Signal and Image Processing, vol. 7883 of *LNCS*, Springer Berlin Heidelberg, 267–278, doi: 10.1007/978-3-642-38294-9_23, 2013.

[4] J. J. van de Gronde, J. B. T. M. Roerdink, Frames for Tensor Field Morphology, in: F. Nielsen, F. Barbaresco (Eds.), Geometric Science of In-

formation, vol. 8085 of *LNCS*, Springer Berlin Heidelberg, 527–534, doi: 10.1007/978-3-642-40020-9_58, 2013.

[5] L. Rittner, R. Lotufo, Diffusion Tensor Imaging Segmentation by Watershed Transform on Tensorial Morphological Gradient, in: Proceedings of SIBGRAPI 2008, IEEE, ISBN 978-0-7695-3358-2, ISSN 1530-1834, 196–203, doi:10.1109/sibgrapi.2008.17, 2008.

[6] A. Alonso-González, S. Valero, J. Chanussot, C. López-Martínez, P. Salembier, Processing Multidimensional SAR and Hyperspectral Images With Binary Partition Tree, Proc. IEEE 101 (3) (2013) 723–747, ISSN 0018-9219, doi:10.1109/jproc.2012.2205209.

[7] S. Velasco-Forero, J. Angulo, Supervised Morphology for Structure Tensor-Valued Images Based on Symmetric Divergence Kernels, in: F. Nielsen, F. Barbaresco (Eds.), Geometric Science of Information, vol. 8085 of *LNCS*, Springer Berlin Heidelberg, 543–550, doi:10.1007/978-3-642-40020-9_60, 2013.

[8] J. Angulo, Supremum/Infimum and Nonlinear Averaging of Positive Definite Symmetric Matrices, in: F. Nielsen, R. Bhatia (Eds.), Matrix Information Geometry, Springer Berlin Heidelberg, 3–33, doi:10.1007/ 978-3-642-30232-9_1, 2013.

[9] B. Burgeth, M. Welk, C. Feddern, J. Weickert, Morphological Operations on Matrix-Valued Images, in: T. Pajdla, J. Matas (Eds.), Computer Vision - ECCV 2004, vol. 3024 of *LNCS*, Springer Berlin Heidelberg, 155–167, doi: 10.1007/978-3-540-24673-2_13, 2004.

[10] B. Burgeth, M. Welk, C. Feddern, J. Weickert, Mathematical Morphology on Tensor Data Using the Loewner Ordering, in: J. Weickert, H. Hagen (Eds.), Visualization and Processing of Tensor Fields, Math. Vis., Springer Berlin Heidelberg, 357–368, doi:10.1007/3-540-31272-2_22, 2006.

[11] B. Burgeth, A. Bruhn, N. Papenberg, M. Welk, J. Weickert, Mathematical morphology for matrix fields induced by the Loewner ordering in higher dimensions, Signal Process. 87 (2) (2007) 277–290, ISSN 01651684, doi: 10.1016/j.sigpro.2005.12.012.

[12] R. Duits, T. C. J. Dela Haije, E. Creusen, A. Ghosh, Morphological and Linear Scale Spaces for Fiber Enhancement in DW-MRI, J. Math. Imaging Vis. 46 (3) (2013) 326–368, ISSN 0924-9907, doi:10.1007/s10851-012-0387-2.

[13] B. Burgeth, S. Didas, L. Florack, J. Weickert, A Generic Approach to the Filtering of Matrix Fields with Singular PDEs, in: F. Sgallari, A. Murli, N. Paragios (Eds.), Scale Space and Variational Methods in Computer Vision, vol. 4485 of *LNCS*, Springer Berlin Heidelberg, 556–567, doi: 10.1007/978-3-540-72823-8_48, 2007.

[14] B. Burgeth, S. Didas, J. Weickert, A General Structure Tensor Concept and Coherence-Enhancing Diffusion Filtering for Matrix Fields, in: D. Laidlaw, J. Weickert (Eds.), Visualization and Processing of Tensor Fields, Math. Vis., Springer Berlin Heidelberg, 305–323, doi:10.1007/978-3-540-88378-4_15, 2009.

[15] G. Birkhoff, Lattice theory, vol. 25 of *American Mathematical Society Colloquium Publications*, American Mathematical Society, 1961.

[16] P. Comon, G. Golub, L.-H. Lim, B. Mourrain, Symmetric Tensors and Symmetric Tensor Rank, SIAM J. Matrix Anal. Appl. 30 (3) (2008) 1254–1279, ISSN 0895-4798, doi:10.1137/060661569.

[17] A. I. Kostrikin, I. I. Manin, Linear algebra and geometry, vol. 1 of *Algebra, Logic and Applications*, Gordon and Breach, ISBN 9056990497, 1997.

[18] N. Bourbaki, Algebra I, Elements of Mathematics, Springer, ISBN 3540642439, 1989.

[19] L. Qi, Eigenvalues of a real supersymmetric tensor, J. Symb. Comput. 40 (6) (2005) 1302–1324, ISSN 07477171, doi:10.1016/j.jsc.2005.05.007.

[20] O. Christensen, Frames and Bases: An Introductory Course, Birkhäuser ; Springer e-books, ISBN 9780817646783, doi:10.1007/978-0-8176-4678-3, 2008.

[21] A. Ben-Israel, T. N. Greville, Generalized inverses: theory and applicatons, Wiley, New York, ISBN 0471065773, 1974.

[22] P. Comon, Tensors versus matrices usefulness and unexpected properties, in: IEEE/SP Workshop on Statistical Signal Processing, IEEE, ISBN 978-1-4244-2709-3, 781–788, doi:10.1109/ssp.2009.5278471, 2009.

[23] T. G. Kolda, B. W. Bader, Tensor Decompositions and Applications, SIAM Review 51 (3) (2009) 455–500, ISSN 0036-1445, doi:10.1137/07070111x.

[24] S. Xu, R. M. Freund, J. Sun, Solution Methodologies for the Smallest Enclosing Circle Problem, Comput. Optim. Appl. 25 (1-3) (2003) 283–292, doi:10.1023/a%253a1022977709811.

[25] U. Köthe, Edge and Junction Detection with an Improved Structure Tensor, in: B. Michaelis, G. Krell (Eds.), Pattern Recognit., vol. 2781 of *LNCS*, Springer Berlin Heidelberg, 25–32, doi:10.1007/978-3-540-45243-0_4, 2003.

[26] K. Chang, L. Qi, T. Zhang, A survey on the spectral theory of nonnegative tensors, Numer. Linear Algebra Appl. 20 (6) (2013) 891–912, doi:10.1002/nla.1902.

[27] M. Breuß, J. Weickert, Highly Accurate PDE-Based Morphology for General Structuring Elements, in: X.-C. Tai, K. Mørken, M. Lysaker, K.-A. Lie (Eds.), Scale Space and Variational Methods in Computer Vision, vol. 5567 of *LNCS*, Springer Berlin Heidelberg, 758–769, doi:10.1007/978-3-642-02256-2_63, 2009.

[28] B. Burgeth, A. Kleefeld, Morphology for Color Images via Loewner Order for Matrix Fields, in: C. L. Luengo Hendriks, G. Borgefors, R. Strand (Eds.), Mathematical Morphology and Its Applications to Signal and Image

Processing, vol. 7883 of *LNCS*, Springer Berlin Heidelberg, 243–254, doi: 10.1007/978-3-642-38294-9_21, 2013.