



Functional Liftings of Vectorial Variational Problems with Laplacian Regularization

Thomas Vogt^(✉) and Jan Lellmann

Institute of Mathematics and Image Computing (MIC), University of Lübeck,
Maria-Goeppert-Str. 3, 23562 Lübeck, Germany
{vogt,lellmann}@mic.uni-luebeck.de

Abstract. We propose a functional lifting-based convex relaxation of variational problems with Laplacian-based second-order regularization. The approach rests on ideas from the calibration method as well as from sublabel-accurate continuous multilabeling approaches, and makes these approaches amenable for variational problems with vectorial data and higher-order regularization, as is common in image processing applications. We motivate the approach in the function space setting and prove that, in the special case of absolute Laplacian regularization, it encompasses the discretization-first sublabel-accurate continuous multilabeling approach as a special case. We present a mathematical connection between the lifted and original functional and discuss possible interpretations of minimizers in the lifted function space. Finally, we exemplarily apply the proposed approach to 2D image registration problems.

Keywords: Variational methods · Curvature regularization · Convex relaxation · Functional lifting · Measure-based regularization

1 Introduction

Let $\Omega \subset \mathbb{R}^d$ and $\Gamma \subset \mathbb{R}^s$ both be bounded sets. In the following, we consider the variational problem of minimizing the functional

$$F(u) = \int_{\Omega} f(x, u(x), \Delta u(x)) dx, \quad (1)$$

that acts on vector-valued functions $u \in C^2(\Omega; \Gamma)$. Convexity of the integrand $f: \Omega \times \Gamma \times \mathbb{R}^s \rightarrow \mathbb{R}$ is only assumed in the last entry, so that $u \mapsto F(u)$ is generally *non-convex*. The Laplacian Δu is understood component-wise and reduces to u'' if the domain Ω is one-dimensional.

Variational problems of this form occur in a wide variety of image processing tasks, including image reconstruction, restoration, and interpolation. Commonly, the integrand is split into data term and regularizer:

$$f(x, z, p) = \rho(x, z) + \eta(p). \quad (2)$$

As an example, in *image registration* (sometimes referred to as large-displacement optical flow), the data term $\rho(x, z) = d(R(x), T(x + z))$ encodes the point-wise distance of a reference image $R: \mathbb{R}^d \rightarrow \mathbb{R}^k$ to a deformed template image $T: \mathbb{R}^d \rightarrow \mathbb{R}^k$ according to a given distance measure $d(\cdot, \cdot)$, such as the squared Euclidean distance $d(a, b) = \frac{1}{2} \|a - b\|_2^2$. While often a suitable convex regularizer η can be found, the highly non-convex nature of ρ renders the search for global minimizers of (1) a difficult problem.

Instead of directly minimizing F using gradient descent or other local solvers, we will aim to *replace* it by a convex functional \mathcal{F} that acts on a higher-dimensional (*lifted*) function space. If the lifting is chosen in such a way that we can construct global minimizers of F from global minimizers of \mathcal{F} , we can find a global solution of the original problem by applying convex solvers to \mathcal{F} . While we cannot claim this property for our choice of lifting, we believe that the mathematical motivation and some of the experimental results show that this approach can be a good basis for future work on global solutions of variational models with higher-order regularization.

Calibrations in Variational Calculus. The lifted functional \mathcal{F} proposed in this work is motivated by previous lifting approaches for *first-order* variational problems of the form

$$\min_u F(u) = \int_{\Omega} f(x, u(x), \nabla u(x)) dx, \tag{3}$$

where F acts on functions $u: \Omega \rightarrow \Gamma$ with $\Omega \subset \mathbb{R}^d$ and scalar range $\Gamma \subset \mathbb{R}$.

The *calibration method* as introduced in [1] gives a globally sufficient optimality condition for functionals of the form (3) with $\Gamma = \mathbb{R}$. Importantly, $f(x, z, p)$ is not required to be convex in (x, z) , but only in p . The method states that u minimizes F if there exists a divergence-free vector field $\phi: \Omega \times \mathbb{R} \rightarrow \mathbb{R}^{d+1}$ (a *calibration*) in a certain admissible set X of vector fields on $\Omega \times \mathbb{R}$ (see below for details), such that

$$F(u) = \int_{\Omega \times \mathbb{R}} \phi \cdot D\mathbf{1}_u, \tag{4}$$

where $\mathbf{1}_u$ is the characteristic function of the subgraph of u in $\Omega \times \mathbb{R}$, $\mathbf{1}_u(x, z) = 1$ if $u(x) > z$ and 0 otherwise, and $D\mathbf{1}_u$ is its distributional derivative. The duality between subgraphs and certain vector fields is also the subject of the broader theory of Cartesian currents [8].

A convex relaxation of the original minimization problem then can be formulated in a higher-dimensional space by considering the functional [5, 19]

$$\mathcal{F}(v) := \sup_{\phi \in X} \int_{\Omega \times \mathbb{R}} \phi \cdot Dv, \tag{5}$$

acting on functions v from the *convex* set

$$\mathcal{C} = \{v: \Omega \times \mathbb{R} \rightarrow [0, 1] : \lim_{z \rightarrow -\infty} v(x, z) = 1, \lim_{z \rightarrow \infty} v(x, z) = 0\}. \tag{6}$$

In both formulations, the set of admissible test functions is

$$X = \{\phi: \Omega \times \mathbb{R} \rightarrow \mathbb{R}^{d+1} : \phi^t(x, z) \geq f^*(x, z, \phi^x(x, z)) \quad (7)$$

$$\text{for every } (x, z) \in \Omega \times \mathbb{R}\}, \quad (8)$$

where $f^*(x, z, p) := \sup_q \langle p, q \rangle - f(x, z, q)$ is the convex conjugate of f with respect to the last variable. In fact, the equality

$$F(u) = \mathcal{F}(\mathbf{1}_u) \quad (9)$$

has been argued to hold for $u \in W^{1,1}(\Omega)$ under suitable assumptions on f [19]. A rigorous proof of the case of $u \in BV(\Omega)$ and $f(x, z, p) = f(z, p)$ (f independent of x), but not necessarily continuous in z , can be found in the recent work [2].

In [17], it is discussed how the choice of discretization influences the results of numerical implementations of this approach. More precisely, motivated by the work [18] from continuous multilabeling techniques, the choice of piecewise linear finite elements on Γ was shown to exhibit so-called *sublabel-accuracy*, which is known to significantly reduce memory requirements.

Vectorial Data. The application of the calibration method to *vectorial* data $\Gamma \subset \mathbb{R}^s$, $s > 1$, is not straightforward, as the concept of subgraphs, which is central to the idea, does not translate easily to higher-dimensional range. While the original sufficient minimization criterion has been successfully translated [16], functional lifting approaches have not been based on this generalization so far. In [20], this approach is considered to be intractable in terms of memory and computational performance.

There are functional lifting approaches for vectorial data with first-order regularization that consider the subgraphs of the components of u [9, 21]. It is not clear how to generalize this approach to nonlinear data $\Gamma \subset \mathcal{M}$, such as a manifold \mathcal{M} , where other functional lifting approaches exist at least for the case of total variation regularization [14].

An approach along the lines of [18] for vectorial data with total variation regularization was proposed in [12]. Even though [17] demonstrated how [18] can be interpreted as a discretized version of the calibration-based lifting, the equivalent approach [12] for vectorial data lacks a fully-continuous formulation as well as a generalization to arbitrary integrands that would demonstrate the exact connection to the calibration method.

Higher-Order Regularization. Another limitation of the calibration method is its limitation to first-order derivatives of u , which leaves out higher-order regularizers such as the Laplacian-based curvature regularizer in image registration [7]. Recently, a functional lifting approach has been successfully applied to second-order regularized image registration problems [15], but the approach was limited to a single regularizer, namely the integral over the 1-norm of the Laplacian (*absolute Laplacian regularization*).

Projection of Lifted Solutions. In the scalar-valued case with first-order regularization, the calibration-based lifting is known to generate minimizers that

can be *projected* to minimizers of the original problem by thresholding [19, Theorem 3.1]. This method is also used for vectorial data with component-wise lifting as in [21]. In the continuous multi-labeling approaches [12, 14, 18], simple averaging is demonstrated to produce useful results even though no theoretical proof is given addressing the accuracy in general. In convex LP relaxation methods, projection (or *rounding*) strategies with provable optimality bounds exist [11] and can be extended to the continuous setting [13]. We demonstrate that rounding is non-trivial in our case, but will leave a thorough investigation to future work.

Contribution. In Sect. 2, we propose a calibration method-like functional lifting approach in the fully-continuous vector-valued setting for functionals that depend in a convex way on Δu . We show that the lifted functional satisfies $\mathcal{F}(\delta_u) \leq F(u)$, where δ_u is the lifted version of a function u and discuss the question of whether the inequality is actually an equality. For the case of absolute Laplacian regularization, we show that our model is a generalization of [15]. In Sect. 2.3, we clarify how convex saddle-point solvers can be applied to our discretized model. Section 3 is concerned with experimental results. We discuss the problem of projection and demonstrate that the model can be applied to image registration problems.

2 A Calibration Method with Vectorial Second-Order Terms

2.1 Continuous Formulation

We propose the following lifted substitute for F :

$$\mathcal{F}(\mathbf{u}) := \sup_{(p,q) \in X} \int_{\Omega} \int_{\Gamma} (\Delta_x p(x, z) + q(x, z)) \, d\mathbf{u}_x(z) dx, \tag{10}$$

acting on functions $\mathbf{u}: \Omega \rightarrow \mathcal{P}(\Gamma)$ with values in the space $\mathcal{P}(\Gamma)$ of Borel probability measures on Γ . This means that, for each $x \in \Omega$ and any measurable set $U \subset \Gamma$, the expression $\mathbf{u}_x(U) \in \mathbb{R}$ can be interpreted as the “confidence” of an assumed underlying function on Ω to take a value inside of U at point x . A function $u: \Omega \rightarrow \Gamma$ can be *lifted* to a function $\mathbf{u}: \Omega \rightarrow \mathcal{P}(\Gamma)$ by defining $\mathbf{u}_x := \delta_{u(x)}$, the Dirac mass at $u(x) \in \Gamma$, for each $x \in \Omega$.

We propose the following set of test functions in the definition of \mathcal{F} :

$$X = \{(p, q) : p \in C_c^2(\Omega \times \Gamma), q \in L^1(\Omega \times \Gamma), \tag{11}$$

$$z \mapsto p(x, z) \text{ concave} \tag{12}$$

$$\text{and } q(x, z) + f^*(x, z, \nabla_z p(x, z)) \leq 0 \tag{13}$$

$$\text{for every } (x, z) \in \Omega \times \Gamma\}, \tag{14}$$

where $f^*(x, z, q) := \sup_{p \in \mathbb{R}^s} \langle q, p \rangle - f(x, z, p)$ is the convex conjugate of f with respect to the last argument.

A thorough analysis of \mathcal{F} requires a careful choice of function spaces in the definition of X as well as a precise definition of the properties of the integrand f and the admissible functions $\mathbf{u}: \Omega \rightarrow \mathcal{P}(\Gamma)$, which we leave to future work. Here, we present a proof that the lifted functional \mathcal{F} bounds the original functional F from below.

Proposition 1. *Let $f: \Omega \times \Gamma \times \mathbb{R}^s \rightarrow \mathbb{R}$ be measurable in the first two, and convex in the third entry, and let $u \in C^2(\Omega; \Gamma)$ be given. Then, for $\mathbf{u}: \Omega \rightarrow \mathcal{P}(\Gamma)$ defined by $\mathbf{u}_x := \delta_{u(x)}$, it holds that*

$$F(u) \geq \mathcal{F}(\mathbf{u}). \tag{15}$$

Proof. Let p, q be any pair of functions satisfying the properties from the definition of X . By the chain rule, we compute

$$\begin{aligned} \Delta_x p(x, u(x)) &= \Delta [p(x, u(x))] - \sum_{i=1}^d \langle \partial_i u(x), D_z^2 p(x, u(x)) \partial_i u(x) \rangle \\ &\quad - 2 \langle \nabla_x \nabla_z p(x, u(x)), \nabla u(x) \rangle - \langle \nabla_z p(x, u(x)), \Delta u(x) \rangle. \end{aligned} \tag{16}$$

Furthermore, the divergence theorem ensures

$$\begin{aligned} - \int_{\Omega} \langle \nabla_x \nabla_z p(x, u(x)), \nabla u(x) \rangle dx &= \int_{\Omega} \langle \nabla_z p(x, u(x)), \Delta u(x) \rangle dx \\ &\quad + \int_{\Omega} \sum_{i=1}^d \langle \partial_i u(x), D_z^2 p(x, u(x)) \partial_i u(x) \rangle dx, \end{aligned} \tag{17}$$

as well as $\int_{\Omega} \Delta [p(x, u(x))] dx = 0$ by the compact support of p . As $p \in C_c^2(\Omega \times \Gamma)$, concavity of $z \mapsto p(x, z)$ implies a negative semi-definite Hessian $D_z^2 p(x, z)$, so that, together with (16)–(17),

$$\int_{\Omega} \Delta_x p(x, u(x)) dx \leq \int_{\Omega} \langle \nabla_z p(x, u(x)), \Delta u(x) \rangle dx. \tag{18}$$

We conclude

$$\mathcal{F}(\mathbf{u}) = \int_{\Omega} \int_{\Gamma} (\Delta_x p(x, z) + q(x, z)) d\mathbf{u}_x(z) dx \tag{19}$$

$$= \int_{\Omega} \Delta_x p(x, u(x)) + q(x, u(x)) dx \tag{20}$$

$$\stackrel{(13)}{\leq} \int_{\Omega} \Delta_x p(x, u(x)) - f^*(x, u(x), \nabla_x p(x, u(x))) dx \tag{21}$$

$$\stackrel{(18)}{\leq} \int_{\Omega} \langle \nabla_z p(x, u(x)), \Delta u(x) \rangle - f^*(x, u(x), \nabla_x p(x, u(x))) dx \tag{22}$$

$$\leq \int_{\Omega} f(x, u(x), \Delta u(x)) dx, \tag{23}$$

where we used the definition of f^* in the last inequality. □

By a standard result from convex analysis, $\langle p, g \rangle - f^*(x, z, g) = f(x, z, p)$ whenever $g \in \partial_p f(x, z, p)$, the subdifferential of f with respect to p . Hence, for equality to hold in (15), we would need to find a function $p \in C_c^2(\Omega \times \Gamma)$ with

$$\nabla_z p(x, u(x)) \in \partial_p f(x, u(x), \Delta u(x)) \tag{24}$$

and associated $q(x, z) := -f^*(x, z, \Delta u(x))$, such that $(p, q) \in X$ or (p, q) can be approximated by functions from X .

Separate Data Term and Regularizer. If the integrand can be decomposed into $f(x, z, p) = \rho(x, z) + \eta(p)$ as in (2), with $\eta \in C^1(\mathbb{R}^s)$ and u sufficiently smooth, the optimal pair (p, q) in the sense of (24) can be explicitly given as

$$p(x, z) := \langle z, \nabla \eta(\Delta u(x)) \rangle, \tag{25}$$

$$q(x, z) := \rho(x, z) - \eta^*(\nabla \eta(\Delta u(x))). \tag{26}$$

A rigorous argument that such p, q exist for any given u could be made by approximating them by compactly supported functions from the admissible set X using suitable cut-off functions on $\Omega \times \Gamma$.

2.2 Connection to the Discretization-First Approach [15]

In [15], data term ρ and regularizer η are lifted independently from each other for the case $\eta = \|\cdot\|_1$. Following the continuous multilabeling approaches in [6, 12, 18], the setting is fully discretized in $\Omega \times \Gamma$ in a first step. Then the lifted data term and regularizer are defined to be the convex hull of a constraint function, which enforces the lifted terms to agree on the Dirac measures δ_u with the original functional applied to the corresponding function u . The data term is taken from [12], while the main contribution concerns the regularizer that now depends on the Laplacian of u .

In this section, we show that our fully-continuous lifting is a generalization of the result from [15] after discretization.

Discretization. In order to formulate the discretization-first lifting approach given in [15], we have to clarify the used discretization.

For the image domain $\Omega \subset \mathbb{R}^d$, discretized using points $X^1, \dots, X^N \in \Omega$ on a rectangular grid, we employ a finite-differences scheme: We assume that, on each grid point X^{i_0} , the discrete Laplacian of $u \in \mathbb{R}^{N,s}$, $u^i \approx u(X^i) \in \mathbb{R}^s$, is defined using the values of u on $m + 1$ grid points X^{i_0}, \dots, X^{i_m} such that

$$(\Delta u)^{i_0} = \sum_{l=1}^m (u^{i_l} - u^{i_0}) \in \mathbb{R}^s. \tag{27}$$

For example, in the case $d = 2$, the popular five-point stencil means $m = 4$ and the X^{i_l} are the neighboring points of X^{i_0} in the rectangular grid. More precisely,

$$\sum_{l=1}^4 (u^{i_l} - u^{i_0}) = [u^{i_1} - 2u^{i_0} + u^{i_2}] + [u^{i_3} - 2u^{i_0} + u^{i_4}]. \tag{28}$$

The range $\Gamma \subset \mathbb{R}^s$ is triangulated into simplices $\Delta_1, \dots, \Delta_M$ with altogether L vertices (or *labels*) $Z^1, \dots, Z^L \in \Gamma$. We write $T := (Z^1 | \dots | Z^L)^T \in \mathbb{R}^{L,s}$, and define the sparse indexing matrices $P^j \in \mathbb{R}^{s+1,L}$ in such a way that the rows of $T_j := P^j T \in \mathbb{R}^{s+1,s}$ are the labels that make up Δ_j .

There exist piecewise linear finite elements $\Phi_k: \Gamma \rightarrow \mathbb{R}$, $k = 1, \dots, L$ satisfying $\Phi_k(t_l) = 1$ if $k = l$, and $\Phi_k(t_l) = 0$ otherwise. In particular, the Φ_k form a partition of unity for Γ , i.e., $\sum_k \Phi_k(z) = 1$ for any $z \in \Gamma$. For a function $p: \Gamma \rightarrow \mathbb{R}$ in the function space spanned by the Φ_k , with a slight abuse of notation, we write $p = (p_1, \dots, p_L)$, where $p_k = p(Z^k)$ so that $p(z) = \sum_k p_k \Phi_k(z)$.

Functional Lifting of the Discretized Absolute Laplacian. Along the lines of classical continuous multilabeling approaches, the absolute Laplacian regularizer is lifted to become the convex hull of the constraint function $\phi: \mathbb{R}^L \rightarrow \mathbb{R} \cup \{+\infty\}$,

$$\phi(p) := \begin{cases} \mu \left\| \sum_{l=1}^m (T_{j_l} \alpha^l - T_{j_0} \alpha^0) \right\|, & \text{if } p = \mu \sum_{l=1}^m (P^{j_l} \alpha^l - P^{j_l} \alpha^0), \\ +\infty, & \text{otherwise,} \end{cases} \quad (29)$$

where $\mu \geq 0$, $\alpha^l \in \Delta_{s+1}^U$ (for Δ_{s+1}^U the unit simplex) and $1 \leq j_l \leq M$ for each $l = 0, \dots, m$. The parameter $\mu \geq 0$ is enforcing positive homogeneity of ϕ which makes sure that the convex conjugate ϕ^* of ϕ is given by the characteristic function $\delta_{\mathcal{K}}$ of a set $\mathcal{K} \subset \mathbb{R}^L$. Namely,

$$\mathcal{K} = \bigcap_{1 \leq j_l \leq M} \{f \in \mathbb{R}^L : \sum_{l=1}^m (f(t^l) - f(t^0)) \leq \left\| \sum_{l=1}^m (t^l - t^0) \right\|, \quad (30)$$

$$\text{for any } \alpha^l \in \Delta_{s+1}^U, l = 0, 1, \dots, m\}, \quad (31)$$

where $t^l := T_{j_l} \alpha^l$ and $f(t^l)$ is the evaluation of the piecewise linear function f defined by the coefficients (f_1, \dots, f_L) (cf. above). The formulation of \mathcal{K} comes with infinitely many constraints so far.

We now show two propositions which give a meaning to this set of constraints for arbitrary dimensions s of the labeling space and an arbitrary choice of norm in the definition of $\eta = \|\cdot\|$. They extend the component-wise (anisotropic) absolute Laplacian result in [15] to the vector-valued case.

Proposition 2. *The set \mathcal{K} can be written as*

$$\mathcal{K} = \{f \in \mathbb{R}^L : f: \Gamma \rightarrow \mathbb{R} \text{ is concave and 1-Lipschitz continuous}\}.$$

Proof. If the piecewise linear function induced by $f \in \mathbb{R}^L$ is concave and 1-Lipschitz continuous, then

$$\frac{1}{m} \sum_{l=1}^m (f(t^l) - f(t^0)) = \left(\frac{1}{m} \sum_{l=1}^m f(t^l) \right) - f(t^0) \leq f \left(\frac{1}{m} \sum_{l=1}^m t^l \right) - f(t^0) \quad (32)$$

$$\leq \left\| \left(\frac{1}{m} \sum_{l=1}^m t^l \right) - t^0 \right\| = \frac{1}{m} \left\| \sum_{l=1}^m (t^l - t^0) \right\|. \quad (33)$$

Hence, $f \in \mathcal{K}$. On the other hand, if $f \in \mathcal{K}$, then we recover Lipschitz continuity by choosing $t^l = t^1$, for any l in (30). For concavity, we first prove mid-point concavity. That is, for any $t^1, t^2 \in \Gamma$, we have

$$\frac{f(t^1)+f(t^2)}{2} \leq f\left(\frac{t^1+t^2}{2}\right) \tag{34}$$

or, equivalently, $[f(t^1) - f(t^0)] + [f(t^2) - f(t^0)] \leq 0$, where $t^0 = \frac{1}{2}(t^1 + t^2)$. This follows from (30) by choosing $t^0 = \frac{1}{2}(t^1 + t^2)$ and $t^l = t^0$ for $l > 2$. With this choice, the right-hand side of the inequality in (30) vanishes and the left-hand side reduces to the desired statement. Now, f is continuous by definition and, for these functions, mid-point concavity is equivalent to concavity. \square

The following theorem is an extension of [15, Theorem 1] to the vector-valued case and is crucial for numerical performance, as it shows that the constraints in Proposition 2 can be reduced to a finite number:

Proposition 3. *The set \mathcal{K} can be expressed using not more than $|\mathcal{E}|$ (nonlinear) constraints, where \mathcal{E} is the set of faces (or edges in the 2D-case) in the triangulation.*

Proof. Usually, Lipschitz continuity of a piecewise linear function requires one constraint on each of the simplices in the triangulation, and thus as many constraints as there are gradients. However, together with concavity, it suffices to enforce a gradient constraint on each of the boundary simplices, of which there are fewer than the number of outer faces in the triangulation. This can be seen by considering the one-dimensional case where Lipschitz constraints on the two outermost pieces of a concave function enforce Lipschitz continuity on the whole domain. Concavity of a function $f : \Gamma \rightarrow \mathbb{R}$ expressed in the basis (Φ_k) is equivalent to its gradient being monotonously decreasing across the common boundary between any neighboring simplices. Together, we need one gradient constraint for each inner, and at most one for each outer face in the triangulation. \square

2.3 Numerical Aspects

For the numerical experiments, we restrict to the special case of integrands $f(x, z, p) = \rho(x, z) + \eta(p)$ as motivated in Sect. 2.1.

Discretization. We base our discretization on the setting in Sect. 2.2. For a function $p: \Gamma \rightarrow \mathbb{R}$ in the function space spanned by the Φ_k , we note that

$$p(z) = \sum_{k=1}^L p_k \Phi_k(z) = \langle A^j z - b^j, P^j p \rangle \text{ whenever } z \in \Delta_j, \tag{35}$$

where A^j and b^j are such that $\alpha = A^j z - b^j \in \Delta_{s+1}^U$ contains the barycentric coordinates of z with respect to Δ_j . More precisely, for $\bar{T}^j := (P^j T | -e)^{-1} \in \mathbb{R}^{s+1, s+1}$ with $e = (1, \dots, 1) \in \mathbb{R}^{s+1}$, we set

$$A^j := \bar{T}^j(1 : \mathbf{s}, :) \in \mathbb{R}^{s, s+1}, \quad b^j := \bar{T}^j(\mathbf{s}+1, :) \in \mathbb{R}^{s+1}. \tag{36}$$

The functions $\mathbf{u}: \Omega \rightarrow \mathcal{P}(I)$ are discretized as $u^{ik} := \int_{\Gamma} \Phi_k(z) d\mathbf{u}_{X^i}(z)$, hence $u \in \mathbb{R}^{N,L}$. Furthermore, whenever $\mathbf{u}_x = \delta_{u(x)}$, the discretization u^i contains the barycentric coordinates of $u(X^i)$ relative to Δ_j . In the context of first-order models, this property is described as sublabeled-accuracy in [12, 17].

Dual Admissibility Constraints. The admissible set X of dual variables is realized by discretizing the conditions (12) and (13).

Concavity (12) of a function $p: \Gamma \rightarrow \mathbb{R}$ expressed in the basis (Φ_k) is equivalent to its gradient being monotonously decreasing across the common boundary between any neighboring simplices. This amounts to

$$\langle g^{j_2} - g^{j_1}, n_{j_1, j_2} \rangle \leq 0, \quad (37)$$

where g^{j_1}, g^{j_2} are the (piecewise constant) gradients $\nabla p(z)$ on two neighboring simplices $\Delta_{j_1}, \Delta_{j_2}$, and $n_{j_1, j_2} \in \mathbb{R}^s$ is the normal of their common boundary pointing from Δ_{j_1} to Δ_{j_2} .

The inequality (13) is discretized using (35) similar to the one-dimensional setting presented in [17]. We denote the dependence of p and q on $X^i \in \Omega$ by a superscript i as in q^i and p^i . Then, for any $j = 1, \dots, M$, we require

$$\sup_{z \in \Delta_j} \langle A^j z - b^j, P^j q^i \rangle - \rho(X^i, z) + \eta^*(g^{ij}) \leq 0 \quad (38)$$

which, for $\rho_j := \rho + \delta_{\Delta_j}$, can be formulated equivalently as

$$\rho_j^*(X^i, (A^j)^T P^j q^i) + \eta^*(g^{ij}) \leq \langle b^j, P^j q \rangle. \quad (39)$$

The fully discretized problem can be expressed in convex-concave saddle point form to which we apply the primal-dual hybrid gradient (PDHG) algorithm [4] with adaptive step sizes from [10]. The epigraph projections for ρ_j^* and η are implemented along the lines of [18, 19].

3 Numerical Results

We implemented the proposed model in Python 3 with NumPy and PyCUDA. The examples were computed on an Intel Core i7 4.00 GHz with 16 GB of memory and an NVIDIA GeForce GTX 1080 Ti with 12 GB of dedicated video memory. The iteration was stopped when the Euclidean norms of the primal and dual residuals [10] fell below $10^{-6} \cdot \sqrt{n}$ where n is the respective number of variables.

Image Registration. We show that the proposed model can be applied to two-dimensional image registration problems (Figs. 1 and 2). We used the sum of squared distances (SSD) data term $\rho(x, z) := \frac{1}{2} \|R(x) - T(x+z)\|_2^2$ and squared Laplacian (curvature) regularization $\eta(p) := \frac{1}{2} \|\cdot\|^2$. The image values $T(x+z)$ were calculated using bilinear interpolation with Neumann boundary conditions. After minimizing the lifted functional, we projected the solution by taking averages over Γ in each image pixel.

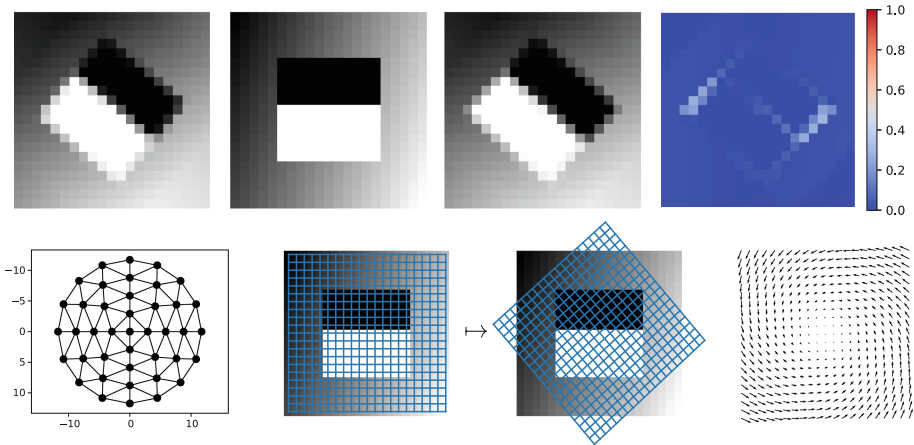


Fig. 1. Application of the proposed higher-order lifting to image registration with SSD data term and squared Laplacian regularization. The method accurately finds a deformation (bottom row, middle and right) that maps the template image (top row, second from left) to the reference image (top row, left), as also visible from the difference image (top row, right). The result (top row, second from right) is almost pixel-accurate, although the range Γ of possible deformation vectors at each point is discretized using only 25 points (second row, left).

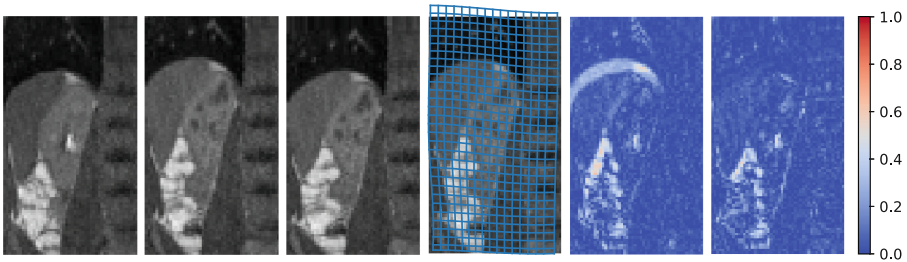


Fig. 2. DCE-MRI data of a human kidney; data courtesy of Jarle Rørvik, Haukeland University Hospital Bergen, Norway; taken from [3]. The deformation (from the left: third and fourth picture) mapping the template (second) to the reference (first) image, computed using our proposed model, is able to significantly reduce the misfit in the left half while fixing the spinal cord at the right edge as can be observed in the difference images from before (fifth) and after (last) registration.

In the first experiment (Fig. 1), the reference image R was synthesized by numerically rotating the template T by 40° . The grid plot of the computed deformation as well as the deformed template are visually very close to the rigid ground-truth deformation (a rotation by 40°). Note that the method obtains almost pixel-accurate results although the range Γ of the deformation is discretized on a disk around the origin, triangulated using only 25 vertices, which is far less than the image resolution.

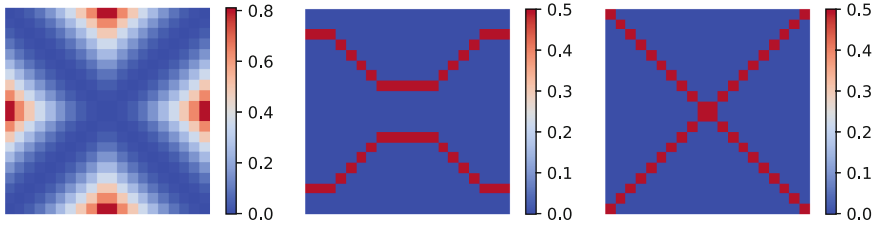


Fig. 3. Minimizers of the lifted functional for the non-convex data term $\rho(x, z) = (|x| - |z|)^2$ (left). With classical first-order total variation-regularized lifting (middle), the result is a composition of two solutions, which can be easily discriminated using thresholding. For the new second-order squared-Laplacian regularized lifting (right), this simple approach fails to separate the two possible (straight line) solutions.

The second experiment (Fig. 2) consists of two coronal slices from a DCE-MRI dataset of a human kidney (data courtesy of Jarle Rørvik, Haukeland University Hospital Bergen, Norway; taken from [3]). The deformation computed using our proposed model is able to significantly reduce the misfit in liver and kidney in the left half while accurately fixing the spinal cord at the right edge.

Projecting the Lifted Solution. In the scalar-valued case with first-order regularization, the minimizers of the calibration-based lifting can be projected to minimizers of the original problem [19, Theorem 3.1]. In our notation, the thresholding technique used there corresponds to mapping \mathbf{u} to

$$u(x) := \inf\{t : \mathbf{u}_x((-\infty, t] \cap \Gamma) > s\}, \tag{40}$$

which is (provably) a global minimizer of the original problem for any $s \in [0, 1)$.

To investigate whether a similar property can hold in our higher-order case, we applied our model with Laplacian regularization $\eta(p) = \frac{1}{2}\|p\|^2$ as well as the calibration method approach with total variation regularization to the data term $\rho(x, z) = (|x| - |z|)^2$ with one-dimensional domain $\Omega = [-1, 1]$ and scalar data $\Gamma = [-1, 1]$ using 20 regularly-spaced discretization points (Fig. 3).

The result from the first-order approach is easily interpretable as a composition of two solutions to the original problem, each of which can be obtained by thresholding (40). In contrast, thresholding applied to the result from the second-order approach yields the two hat functions $v_1(x) = |x|$ and $v_2(x) = -|x|$, neither of which minimizes the original functional. Instead, the solution turns out to be of the form $\mathbf{u} = \frac{1}{2}\delta_{u_1} + \frac{1}{2}\delta_{u_2}$, where u_1 and u_2 are in fact global minimizers of the original problem: namely, the straight lines $u_1(x) = x$ and $u_2(x) = -x$.

4 Conclusion

In this work we presented a novel fully-continuous functional lifting approach for non-convex variational problems that involve Laplacian second-order terms

and vectorial data, with the aim to ultimately provide sufficient optimality conditions and find global solutions despite the non-convexity. First experiments indicate that the method can produce subpixel-accurate solutions for the non-convex image registration problem. We argued that more involved projection strategies than in the classical calibration approach will be needed for obtaining a good (approximate) solution of the original problem from a solution of the lifted problem. Another interesting direction for future work is the generalization to functionals that involve arbitrary second- or higher-order terms.

Acknowledgments. The authors acknowledge support through DFG grant LE 4064/1-1 “Functional Lifting 2.0: Efficient Convexifications for Imaging and Vision” and NVIDIA Corporation.

References

1. Alberti, G., Bouchitté, G., Dal Maso, G.: The calibration method for the mumford-shah functional and free-discontinuity problems. *Calc. Var. Partial Differ. Equ.* **16**(3), 299–333 (2003)
2. Bouchitté, G., Fragalà, I.: A duality theory for non-convex problems in the calculus of variations. *Arch. Rational. Mech. Anal.* **229**(1), 361–415 (2018)
3. Brehmer, K., Wacker, B., Modersitzki, J.: A novel similarity measure for image sequences. In: Klein, S., Staring, M., Durrleman, S., Sommer, S. (eds.) *WBIR 2018*. LNCS, vol. 10883, pp. 47–56. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-92258-4_5
4. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vis.* **40**(1), 120–145 (2011)
5. Chambolle, A.: Convex representation for lower semicontinuous envelopes of functionals in L^1 . *J. Convex Anal.* **8**(1), 149–170 (2001)
6. Chambolle, A., Cremers, D., Pock, T.: A convex approach to minimal partitions. *SIAM J. Imaging Sci.* **5**(4), 1113–1158 (2012)
7. Fischer, B., Modersitzki, J.: Curvature based image registration. *J. Math. Imaging Vis.* **18**(1), 81–85 (2003)
8. Giaquinta, M., Modica, G., Souček, J.: Cartesian currents in the calculus of variations I. *Cartesian Currents*, vol. 37. Springer, Heidelberg (1998)
9. Goldluecke, B., Strelakovsky, E., Cremers, D.: Tight convex relaxations for vector-valued labeling. *SIAM J. Imaging Sci.* **6**(3), 1626–1664 (2013)
10. Goldstein, T., Esser, E., Baraniuk, R.: Adaptive primal dual optimization for image processing and learning. In: *Proceedings 6th NIPS Workshop Optimize Machine Learning* (2013)
11. Kleinberg, J., Tardos, E.: Approximation algorithms for classification problems with pairwise relationships: metric labeling and markov random fields. *J. ACM (JACM)* **49**(5), 616–639 (2002)
12. Laude, E., Möllenhoff, T., Moeller, M., Lellmann, J., Cremers, D.: Sublabel-accurate convex relaxation of vectorial multilabel energies. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) *ECCV 2016*. LNCS, vol. 9905, pp. 614–627. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46448-0_37
13. Lellmann, J., Lenzen, F., Schnörr, C.: Optimality bounds for a variational relaxation of the image partitioning problem. *J. Math. Imaging Vis.* **47**(3), 239–257 (2013)

14. Lellmann, J., Strelakovski, E., Koetter, S., Cremers, D.: Total variation regularization for functions with values in a manifold. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2944–2951 (2013)
15. Loewenhauser, B., Lellmann, J.: Functional lifting for variational problems with higher-order regularization. In: Tai, X.-C., Bae, E., Lysaker, M. (eds.) IVLOPDE 2016. MV, pp. 101–120. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-91274-5_5
16. Mora, M.G.: The calibration method for free-discontinuity problems on vector-valued maps. *J. Convex Anal.* **9**(1), 1–29 (2002)
17. Möllenhoff, T., Cremers, D.: Sublabel-accurate discretization of nonconvex free-discontinuity problems. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV), October 2017
18. Möllenhoff, T., Laude, E., Moeller, M., Lellmann, J., Cremers, D.: Sublabel-accurate relaxation of nonconvex energies. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3948–3956 (2016)
19. Pock, T., Cremers, D., Bischof, H., Chambolle, A.: Global solutions of variational models with convex regularization. *SIAM J. Imaging Sci.* **3**(4), 1122–1145 (2010)
20. Strelakovski, E., Chambolle, A., Cremers, D.: A convex representation for the vectorial mumford-shah functional. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1712–1719. IEEE (2012)
21. Strelakovski, E., Chambolle, A., Cremers, D.: Convex relaxation of vectorial problems with coupled regularization. *SIAM J. Imaging Sci.* **7**(1), 294–336 (2014)