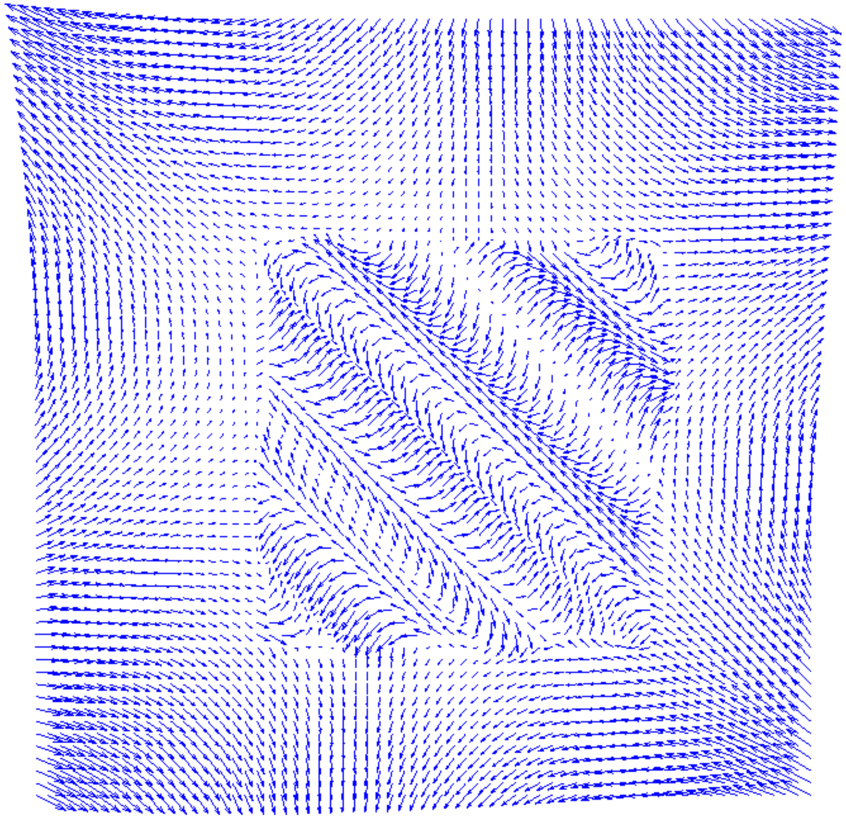
On TV Regularization of Image Flows and MRF-Inference by DC-Programming

Jörg Kappes, Christoph Schnörr, Jing Yuan
University of Heidelberg

"Graph Cuts and Related Discrete
or Continuous Optimization Problems"
IPAM, Feb 25-29, 2008

TV regularization of image flows

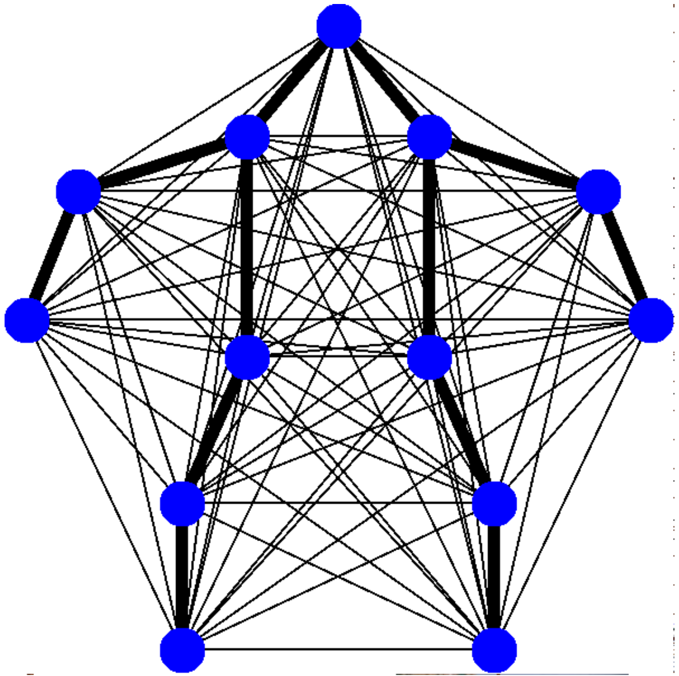


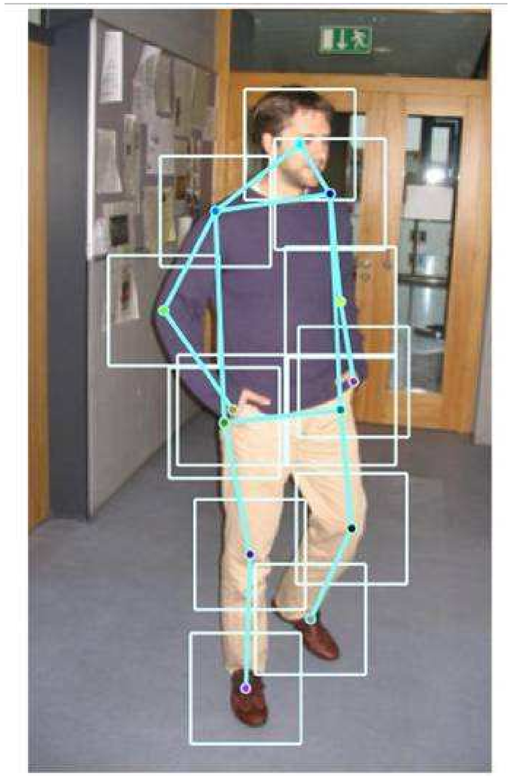
- Contextual decisions by convex programming
- ROF-model, piecewise constant scalar fields $\longrightarrow$ piecewise harmonic vector fields
- Length-regularization (Mumford-Shah)

Related work:

Aujol, Chambolle, Chan, Darbon, Esedoglu, Fatemi, Gilboa, Goldfarb, Meyer, Nikolova, Osher, Rudin, Shen, Tai, Vese, Yin, ...

MRF-Inference by DC-Programming





- $5 \dots 25$ vertices
- $20 \dots 100$ states

Fairly different from grid-graphs.

Related work: *Ravikumar&Lafferty*, *audience*.

TV regularization of image flows

orthog. decomposition

$$L^2(\Omega) = H_c \oplus \operatorname{div} P$$

$$L^2(\Omega)^2 = \mathcal{H} \oplus \nabla H_0^1(\Omega) \oplus \nabla^\perp H_0^1(\Omega)$$

convex decomposition

$$L^2(\Omega) = U \oplus_\Pi \operatorname{div} C_\lambda$$

$$L^2(\Omega)^2 = U_{\mathcal{H}} \oplus_\Pi \left\langle (\nabla, \nabla^\perp) C_\lambda \right\rangle$$

ROF model, convex image decomposition

$$\frac{1}{2}\|u-d\|_{\Omega}^{2}+\lambda \mathrm{TV}(u)$$

$$\mathrm{TV}(u)=\sup_{\|p\|_\infty\leq 1}\langle u, \operatorname{div} p\rangle_\Omega\ , \quad p|_{\partial\Omega}=0$$

Dual formulation

$$\min_{\|p\|_\infty\leq 1}\Big\{\|\lambda \operatorname{div} p-d\|_\Omega^2\Big\}\ , \quad d=u+\lambda \operatorname{div} p\ , \quad p|_{\partial\Omega}=0$$

Orthog. projector

$$\Pi\colon\, L^2(\Omega)\to \operatorname{div}\, C_\lambda$$

$$C_\lambda=\big\{(p_1,p_2)^\top \;\big|\; \|p\|_\infty\leq\lambda\big\}$$

ROF model, convex image decomposition

$$d = u + v \ , \qquad L^2(\Omega) = U \oplus_{\Pi} \operatorname{div} C_{\lambda}$$

Orthogonal image decomposition

Components

$$d = c + \operatorname{div} p \,, \qquad p|_{\partial\Omega} = 0$$
$$= \frac{1}{|\Omega|}\langle 1, d\rangle_\Omega + \operatorname{div} p$$

Orthogonality

$$\langle c, \operatorname{div} p\rangle_\Omega = 0 \,, \qquad L^2(\Omega) = H_c \oplus \operatorname{div} P$$

Notation

Scalar curl operator for vector fields v

$$\operatorname{curl} v = \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2}$$

Vector-valued curl operator for scalar fields p

$$\nabla^{\perp} p = \left(\frac{\partial p}{\partial x_2} \, , \; - \frac{\partial p}{\partial x_1} \right)^{\top}$$

Orthogonal vector field decomposition

Basic subspaces

$$H(\mathrm{div};\Omega) = \left\{ v \in L^2(\Omega)^2 \;\middle|\; \operatorname{div} v \in L^2(\Omega) \right\} ,$$
$$H(\mathrm{curl};\Omega) = \left\{ v \in L^2(\Omega)^2 \;\middle|\; \operatorname{curl} v \in L^2(\Omega) \right\} .$$

Standard orthogonal decompositions

$$L^2(\Omega)^2 = \nabla H^1(\Omega) \oplus \nabla^\perp H_0^1(\Omega) ,$$
$$L^2(\Omega)^2 = \nabla H_0^1(\Omega) \oplus \nabla^\perp H^1(\Omega) .$$

Gradients and curls of $H^1(\Omega)$

$$\nabla H^1(\Omega) = \left\{ v \in H(\mathrm{curl};\Omega) \;\middle|\; \operatorname{curl} v = 0 \right\} ,$$
$$\nabla^\perp H^1(\Omega) = \left\{ v \in H(\mathrm{div};\Omega) \;\middle|\; \operatorname{div} v = 0 \right\} .$$

Orthogonal vector field decomposition

Refined decomposition

$$L^2(\Omega)^2 = \mathcal{H} \oplus \nabla H_0^1(\Omega) \oplus \nabla^{\perp} H_0^1(\Omega)$$

Harmonic vector fields

$$\begin{aligned} \mathcal{H} &= \Big\{ v \in H(\mathrm{div};\Omega) \cap H(\mathrm{curl};\Omega) \;\Big|\; \mathrm{div}\, v = \mathrm{curl}\, v = 0 \Big\} \;, \\ &= \nabla^{\perp} H^1(\Omega) \cap \nabla H^1(\Omega) \;. \end{aligned}$$

Convex decomposition of vector fields

Components

$$\begin{aligned} d &= u + v \ , \\ &= u + \nabla\psi + \nabla^{\perp}\phi \ . \end{aligned}$$

Orthog. projector

$$\Pi \colon L^2(\Omega)^2 \to S_\lambda$$

$$S_\lambda = \left\{ \nabla\psi + \nabla^{\perp}\phi \;\middle|\; (\psi,\phi) \in C_\lambda \cap H^1_0(\Omega)^2 \right\}$$

Decomposition

$$L^2(\Omega)^2 = U_{\mathcal{H}} \ \oplus_{\Pi} \ \left\langle (\nabla, \nabla^{\perp}), C_\lambda \right\rangle$$

Denoising

Regularizer

$$R(u) = \sup_{s \in S} \langle s, u \rangle_\Omega \ ,$$

$$= \sup_{\|(\psi,\phi)\|_\infty \leq 1} \langle \nabla\psi + \nabla^\perp \phi, u \rangle_\Omega \ ,$$

$$= \sup_{\|(\psi,\phi)\|_\infty \leq 1} \Big\{ - \langle \psi, \operatorname{div} u \rangle_\Omega - \langle \phi, \operatorname{curl} u \rangle_\Omega \Big\} \ ,$$

$$= \int_\Omega \sqrt{(\operatorname{div} u)^2 + (\operatorname{curl} u)^2} dx \ .$$

Differs from

$$\mathrm{TV}(u) = \int_\Omega \sqrt{|\nabla u_1|^2 + |\nabla u_2|^2}$$

Motion boundaries

Assume $u = u|_{\Omega_1} + u|_{\Omega_2}$ to be piecewise harmonic.

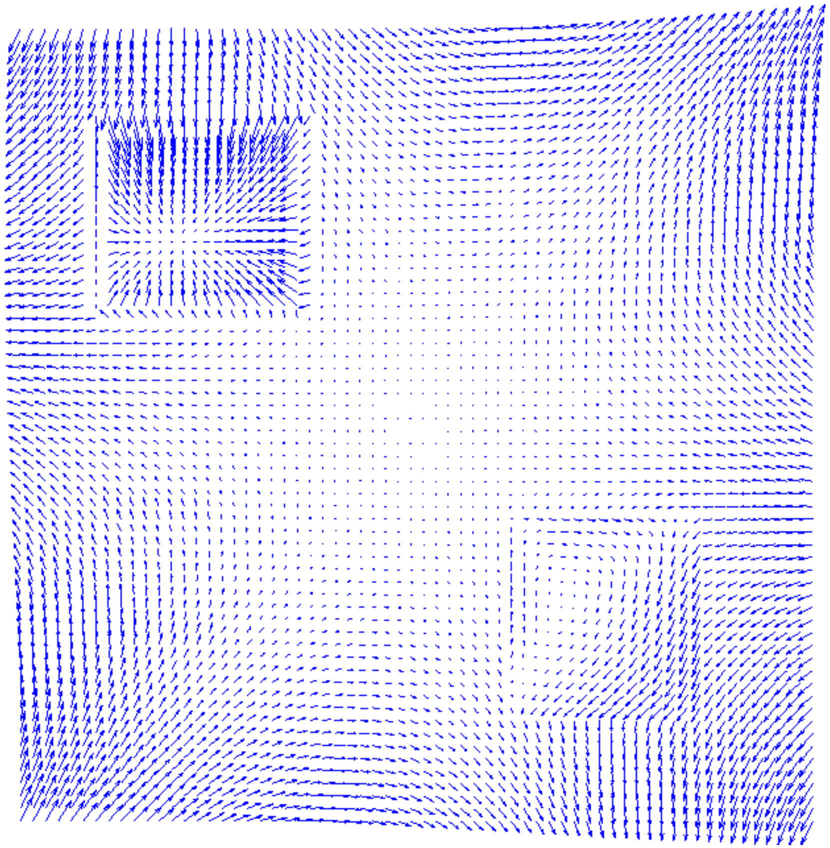
$$\begin{aligned} R(u) &= \sup_{\|(\psi,\phi)\|_\infty \leq 1} \langle \nabla\psi + \nabla^\perp \phi, u\rangle_\Omega \ , \\ &= \sup_{\|(\psi,\phi)\|_\infty \leq 1} \Big\{ -\langle \psi, \operatorname{div} u\rangle_{\Omega_1} + \int_{\partial(\Omega_1,\Omega_2)} \langle u_1 - u_2, n\rangle \psi \, ds \\ &\qquad\qquad + \langle \phi, \operatorname{curl} u\rangle_{\Omega_2} + \int_{\partial(\Omega_1,\Omega_2)} \langle u_1 - u_2, t\rangle \phi \, ds \Big\} \ , \\ &= \int_{\partial(\Omega_1,\Omega_2)} \|u_1 - u_2\| ds \ . \end{aligned}$$

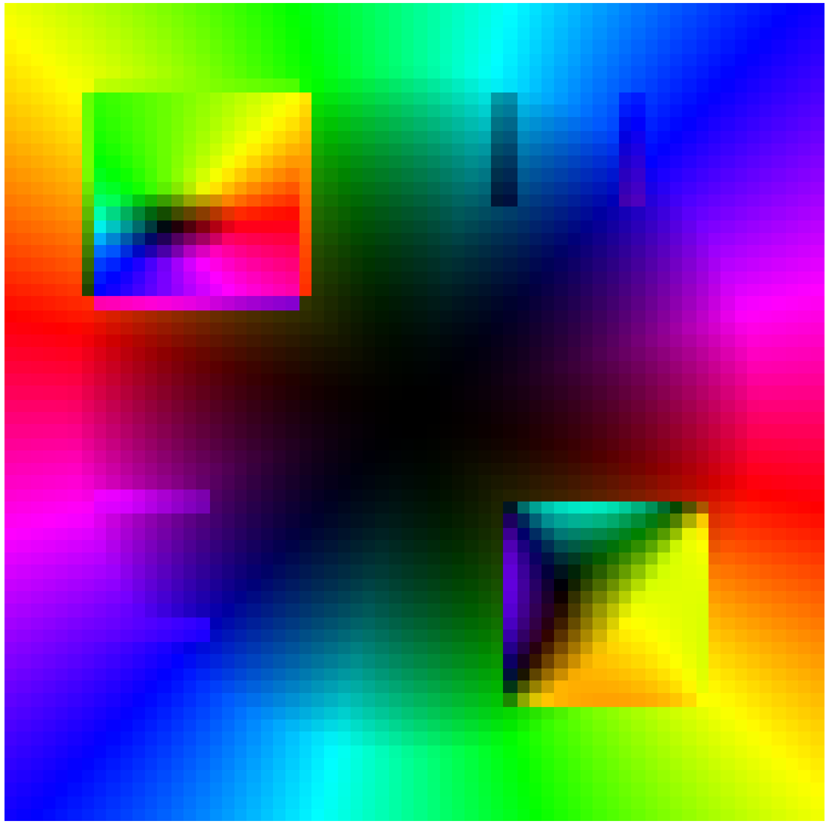
Generalizes the term

$$(u_1 - u_2) \text{ length}\big(\partial(\Omega_1, \Omega_2)\big)$$

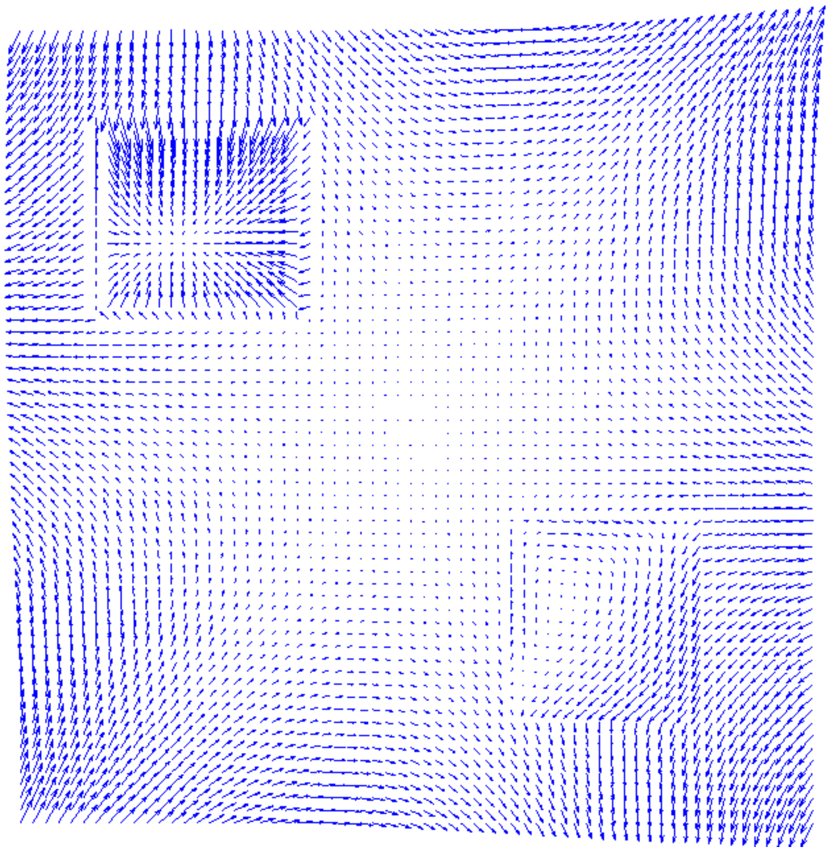
valid for *piecewise constant scalar* fields.

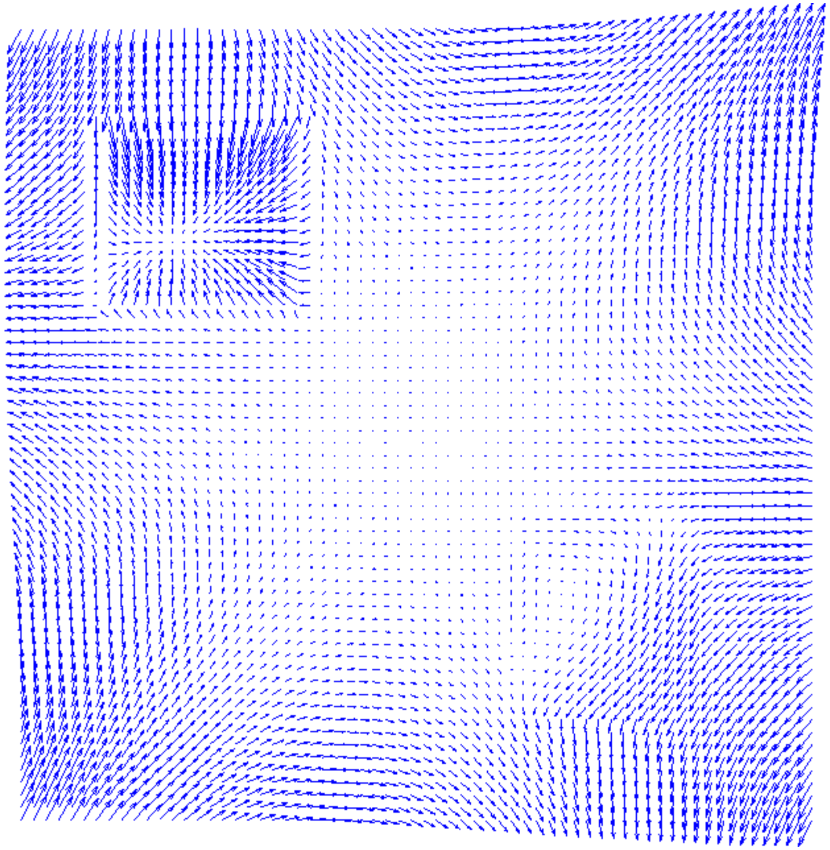
Input data



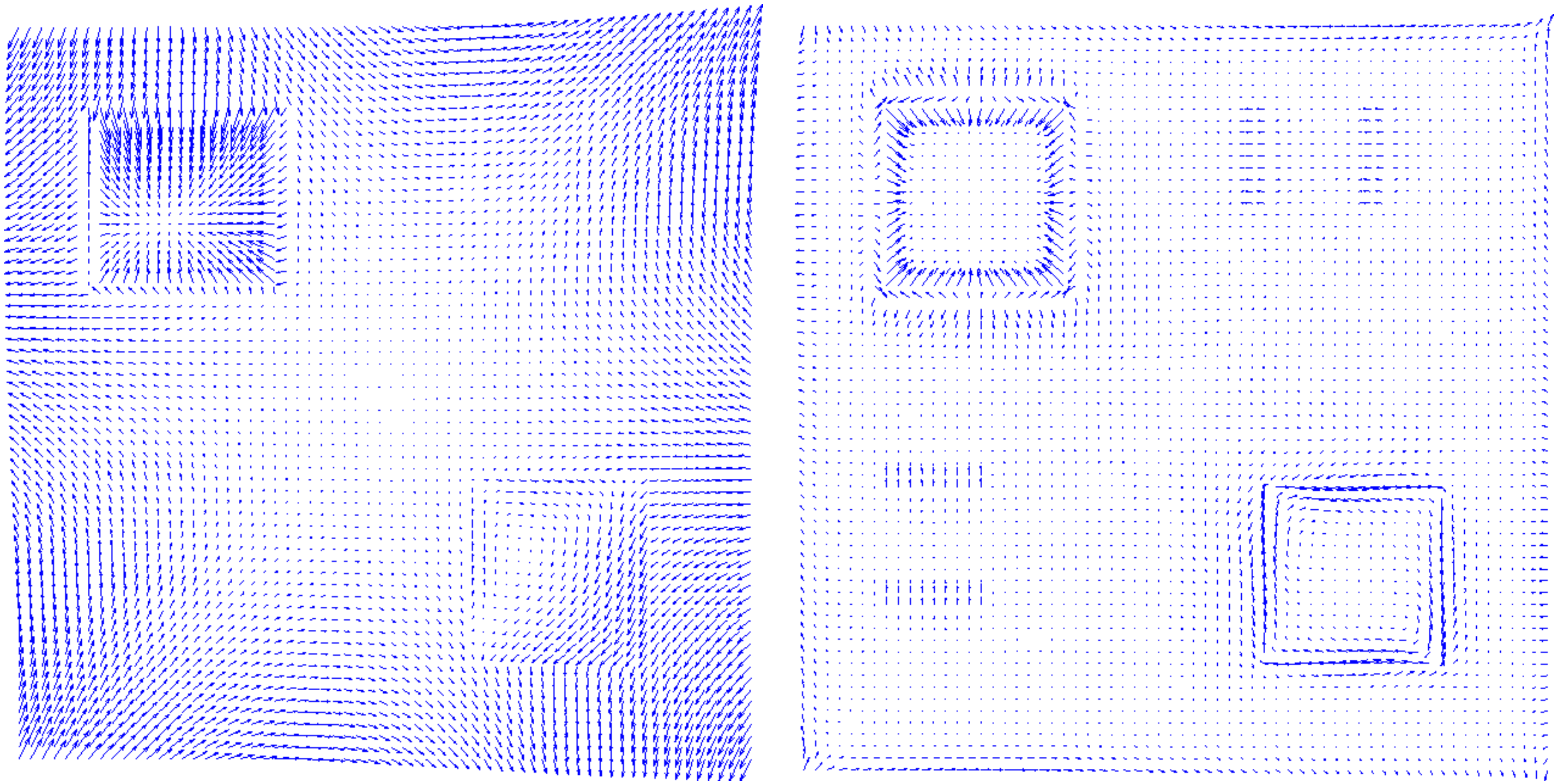


Standard TV-term, small λ, u-component

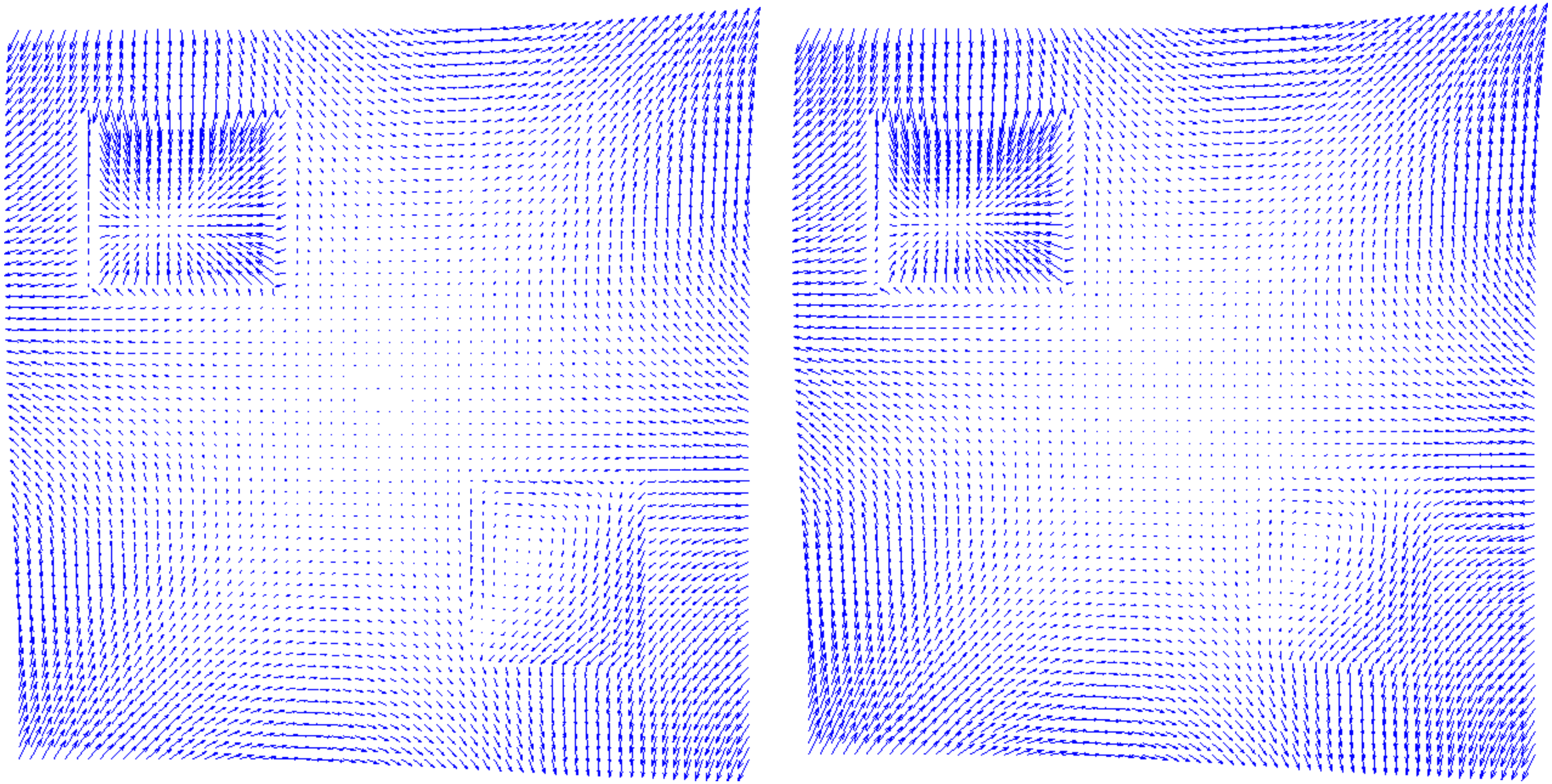




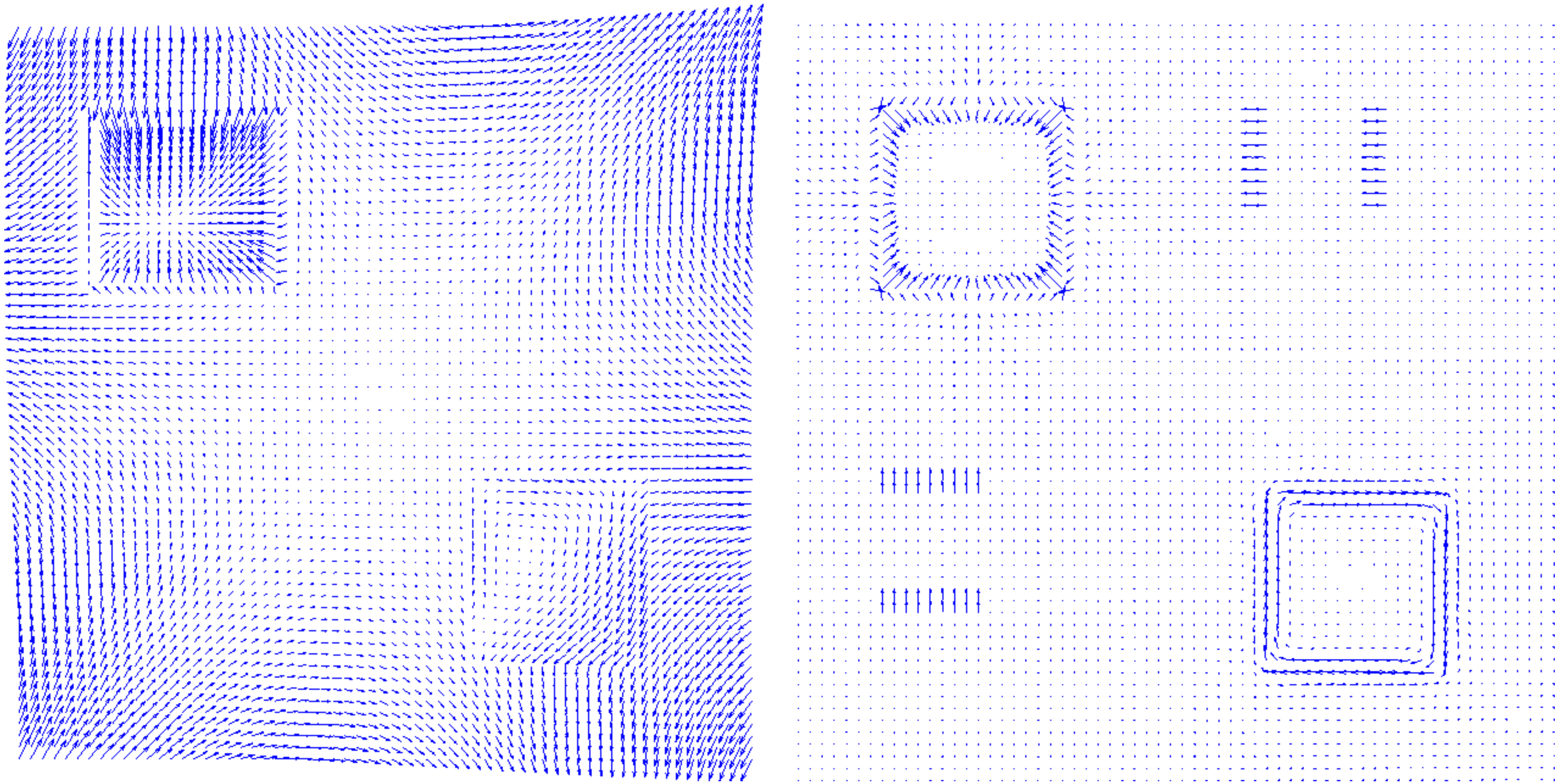
Standard TV-term, small λ, v-component



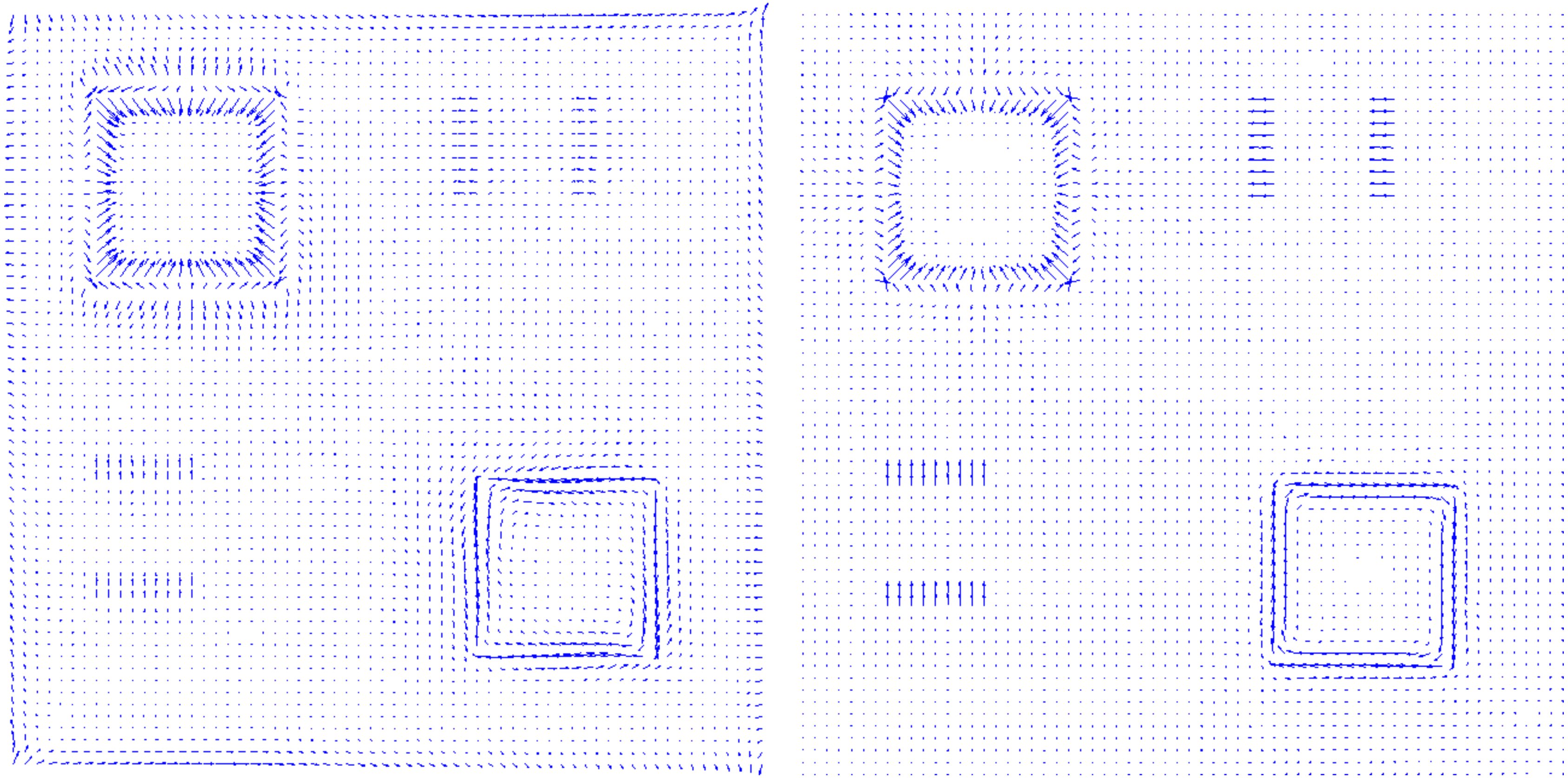
Alternativ TV-term, small λ, u-component



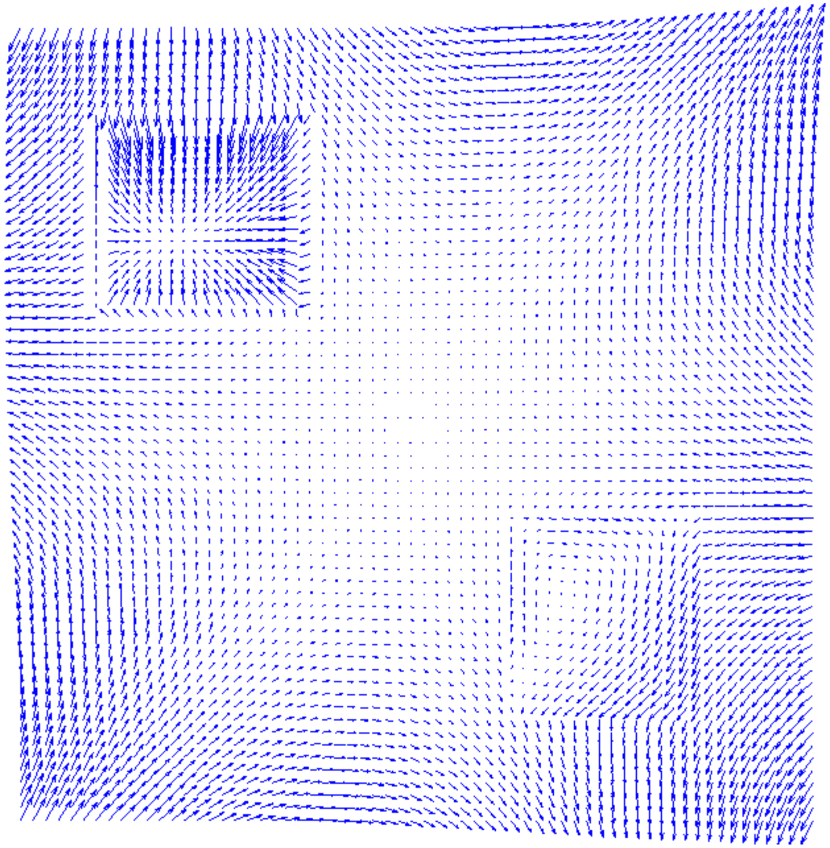
Alternativ TV-term, small λ, v-component

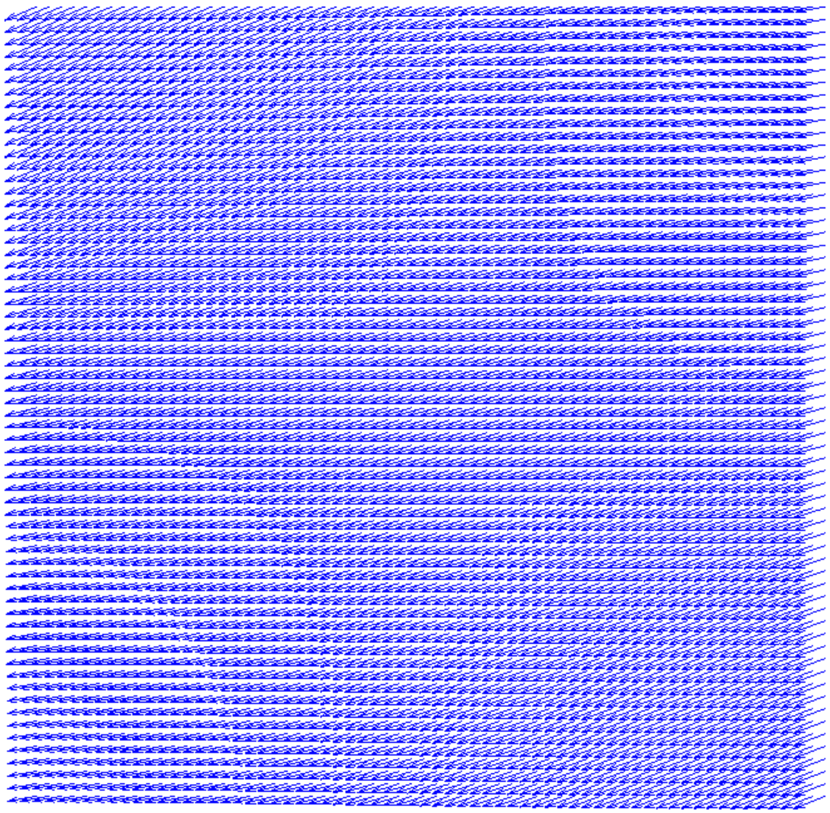


v-components

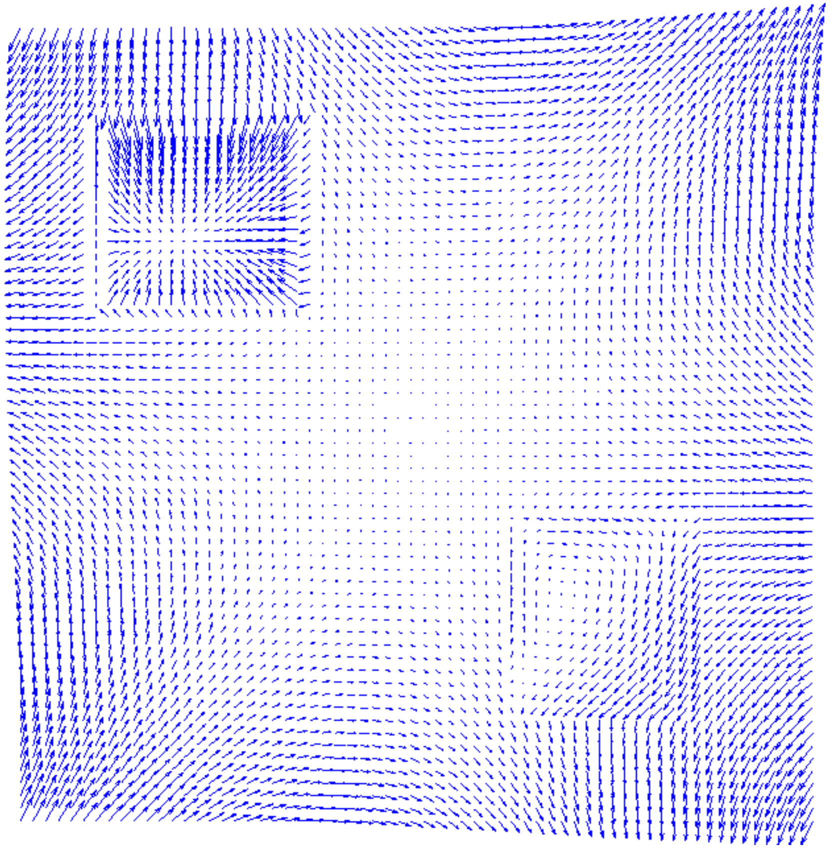


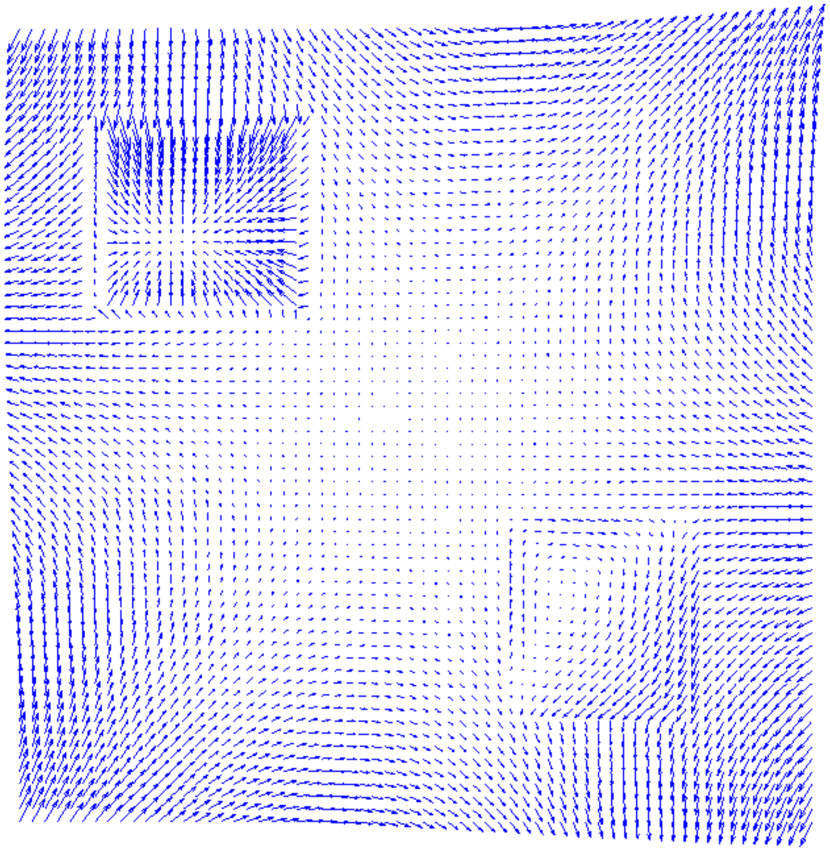
Standard TV-term, large λ, u-component



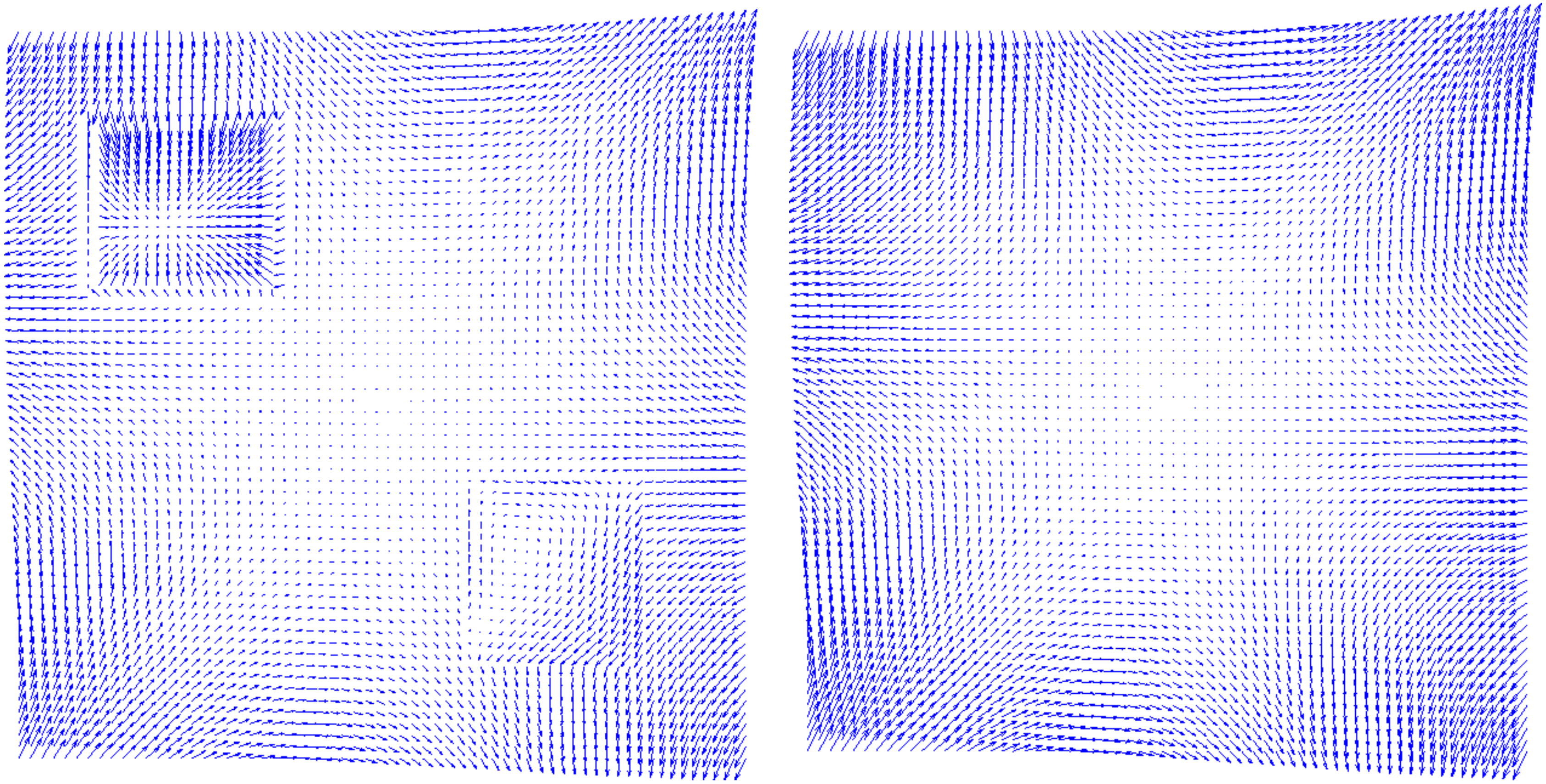


Standard TV-term, large λ, v-component

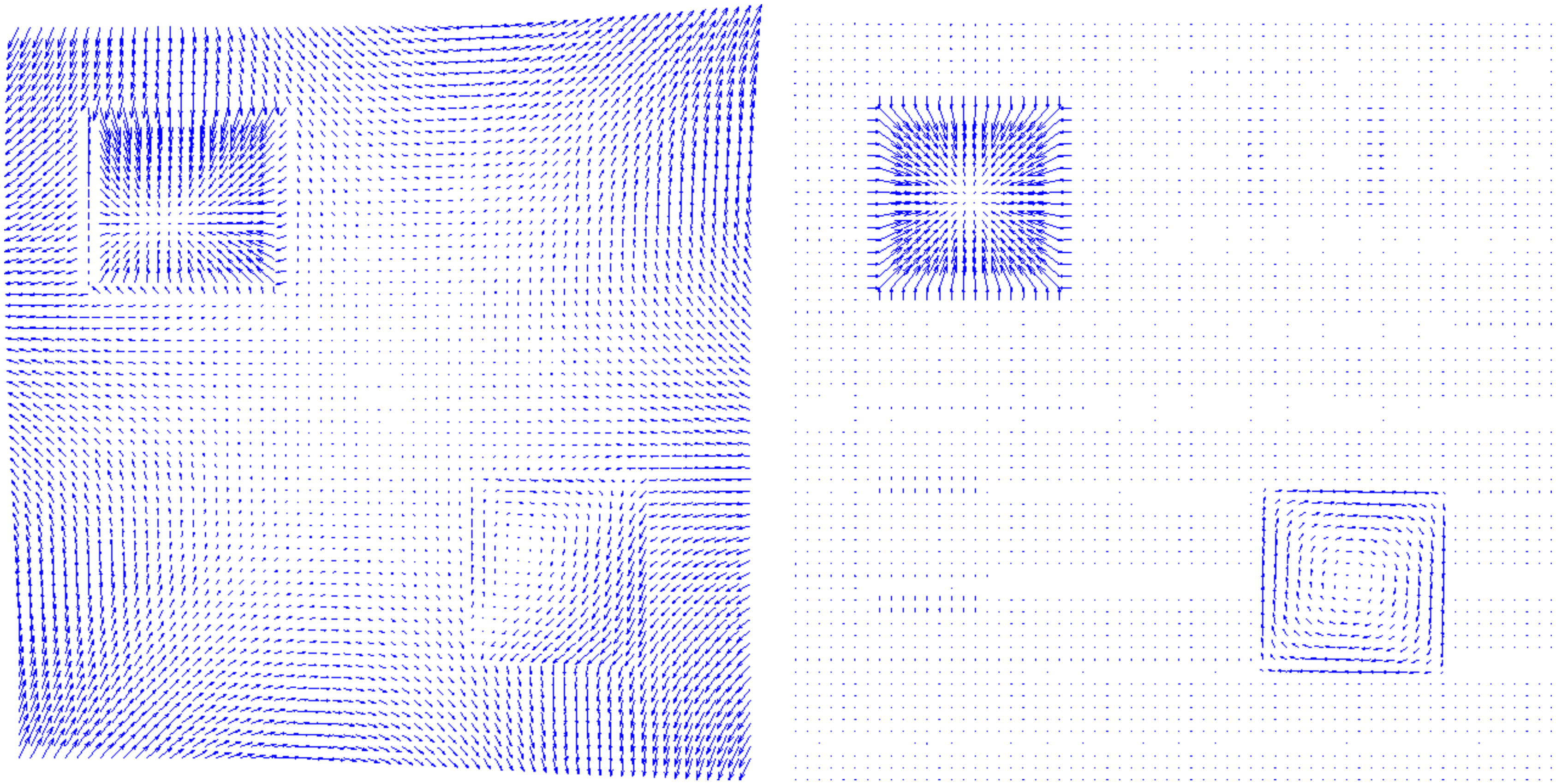




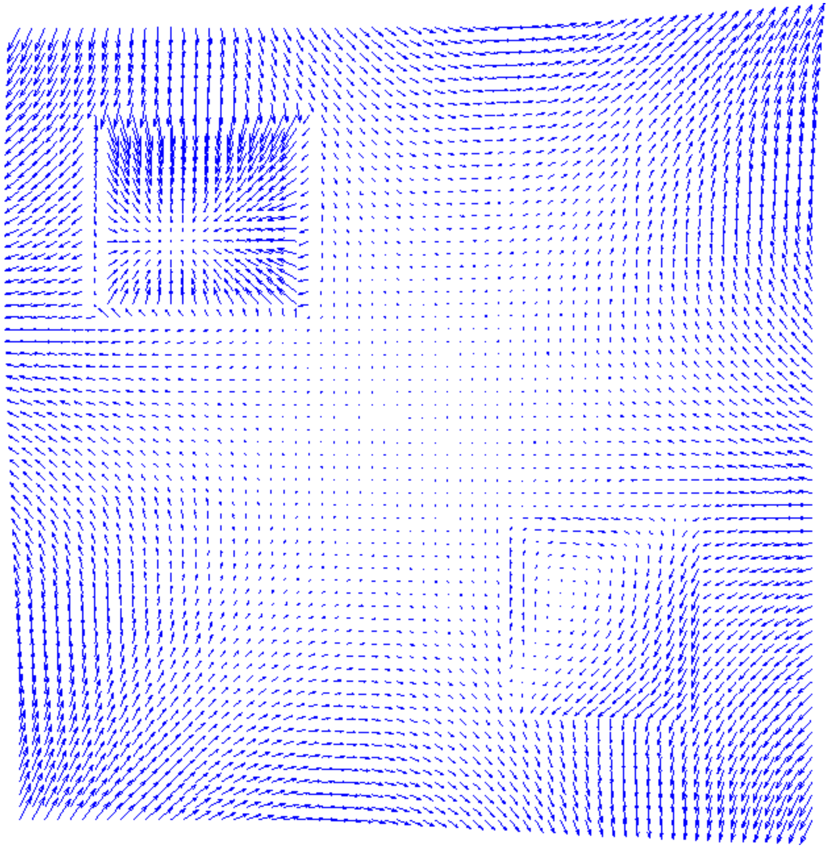
Alternative TV-term, large λ, u-component

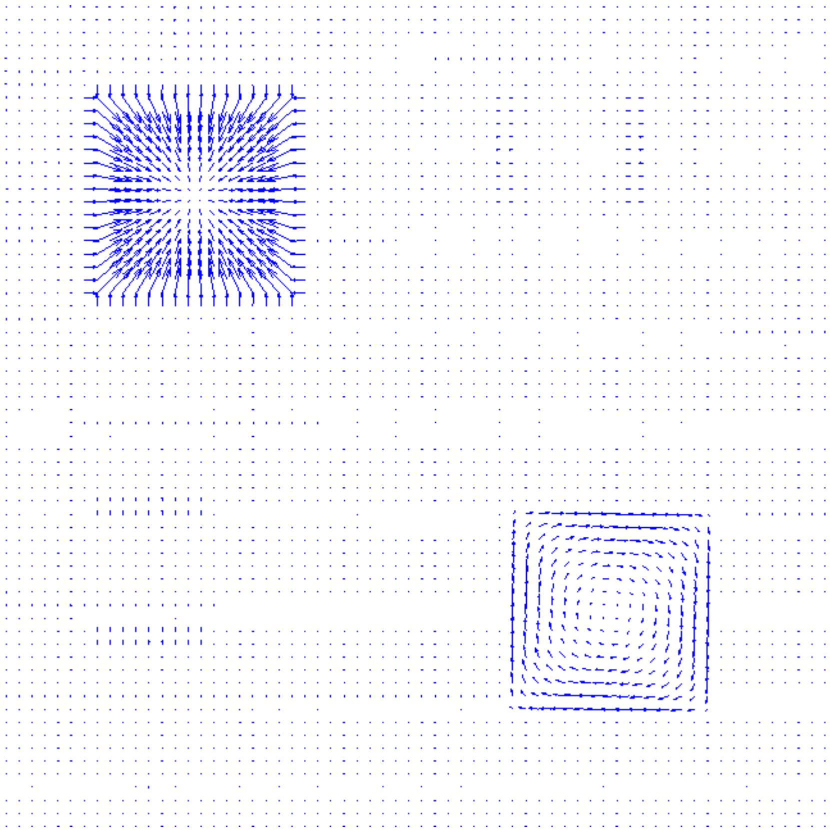


Alternative TV-term, large λ, v-component

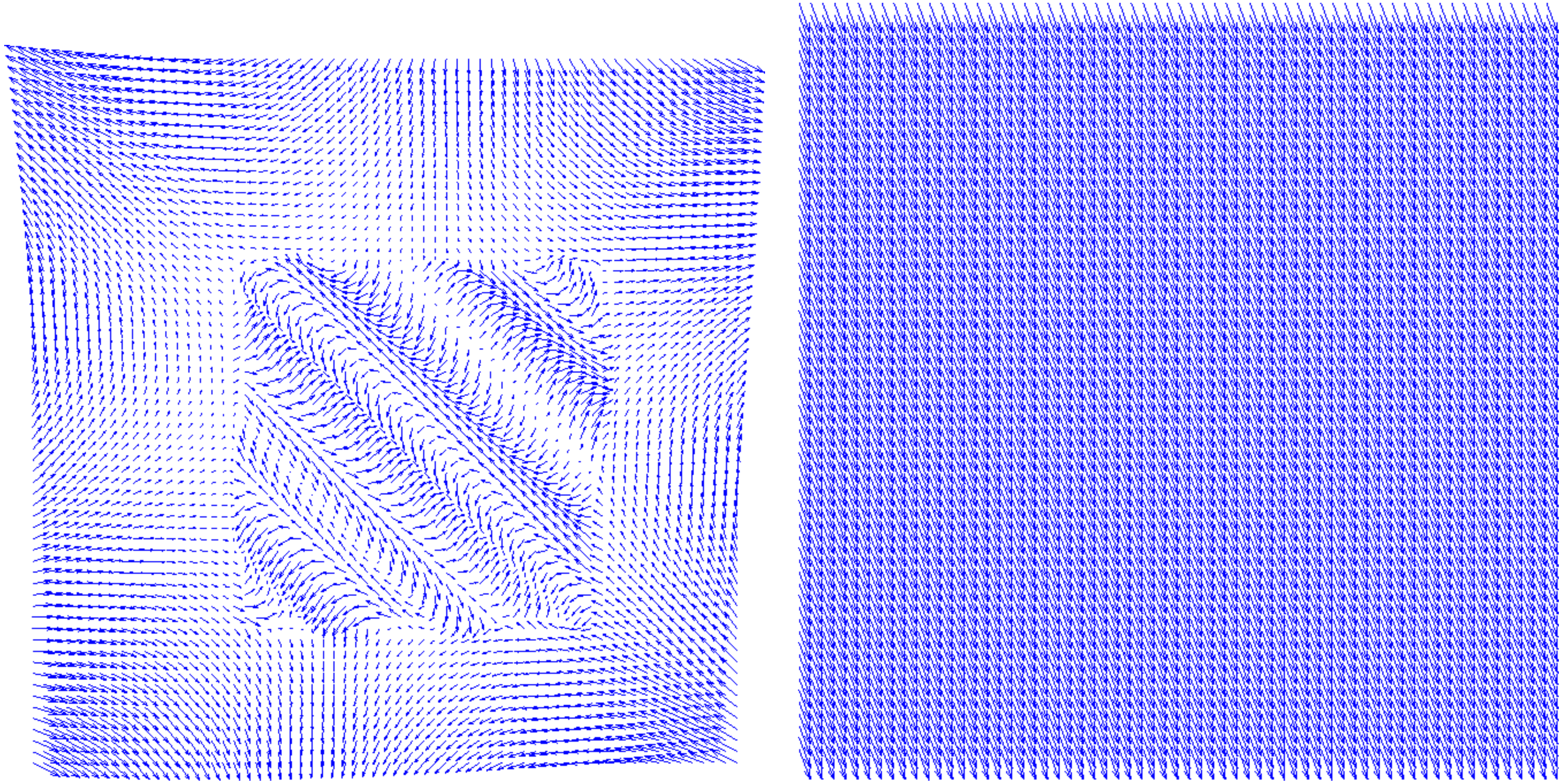


v-components

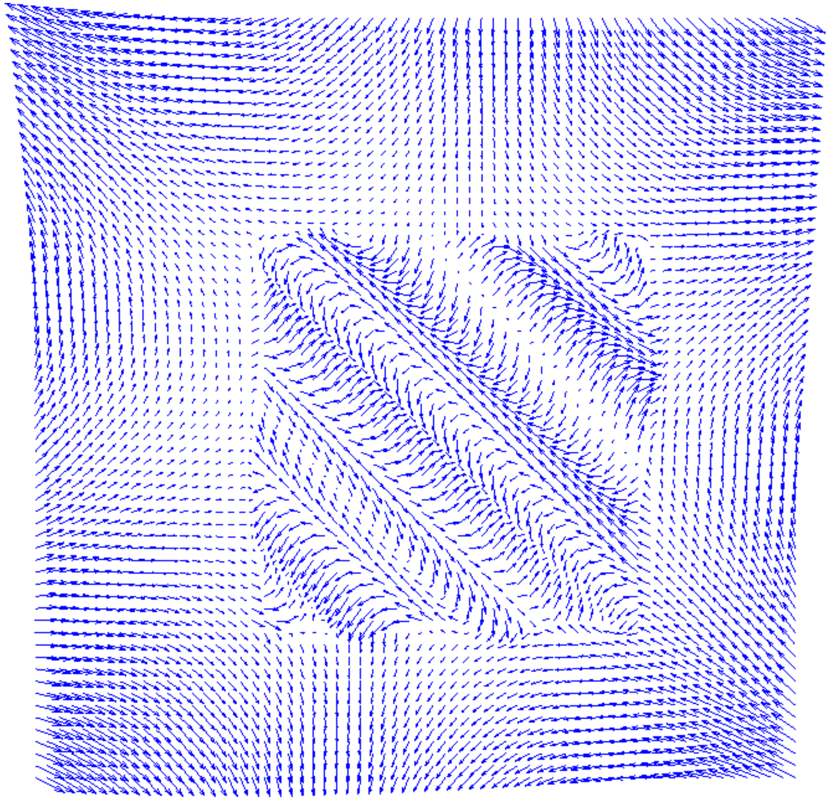


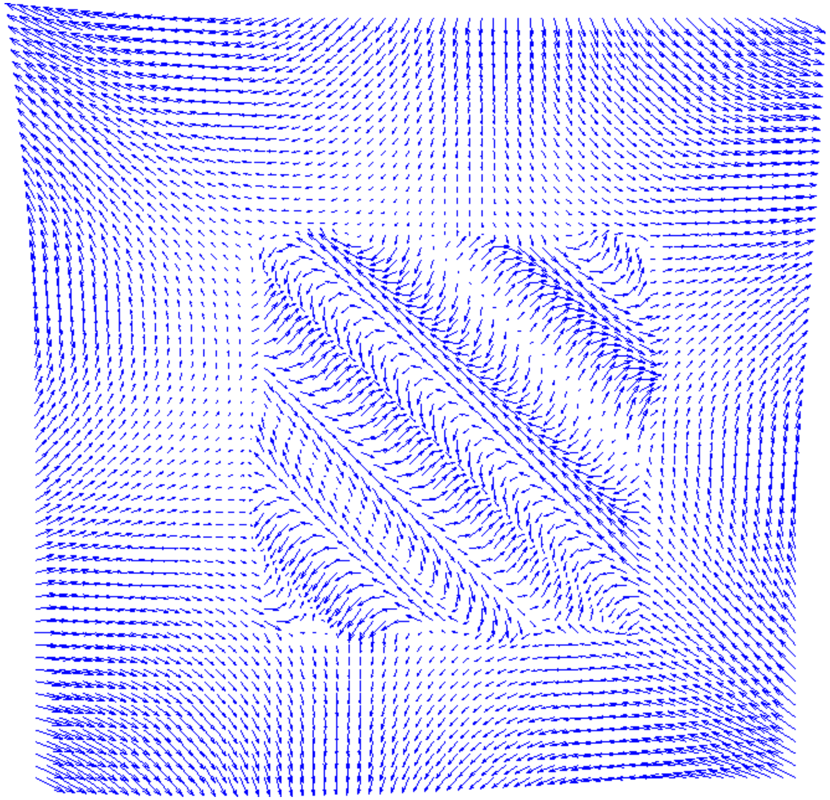


Standard TV-term, u-component

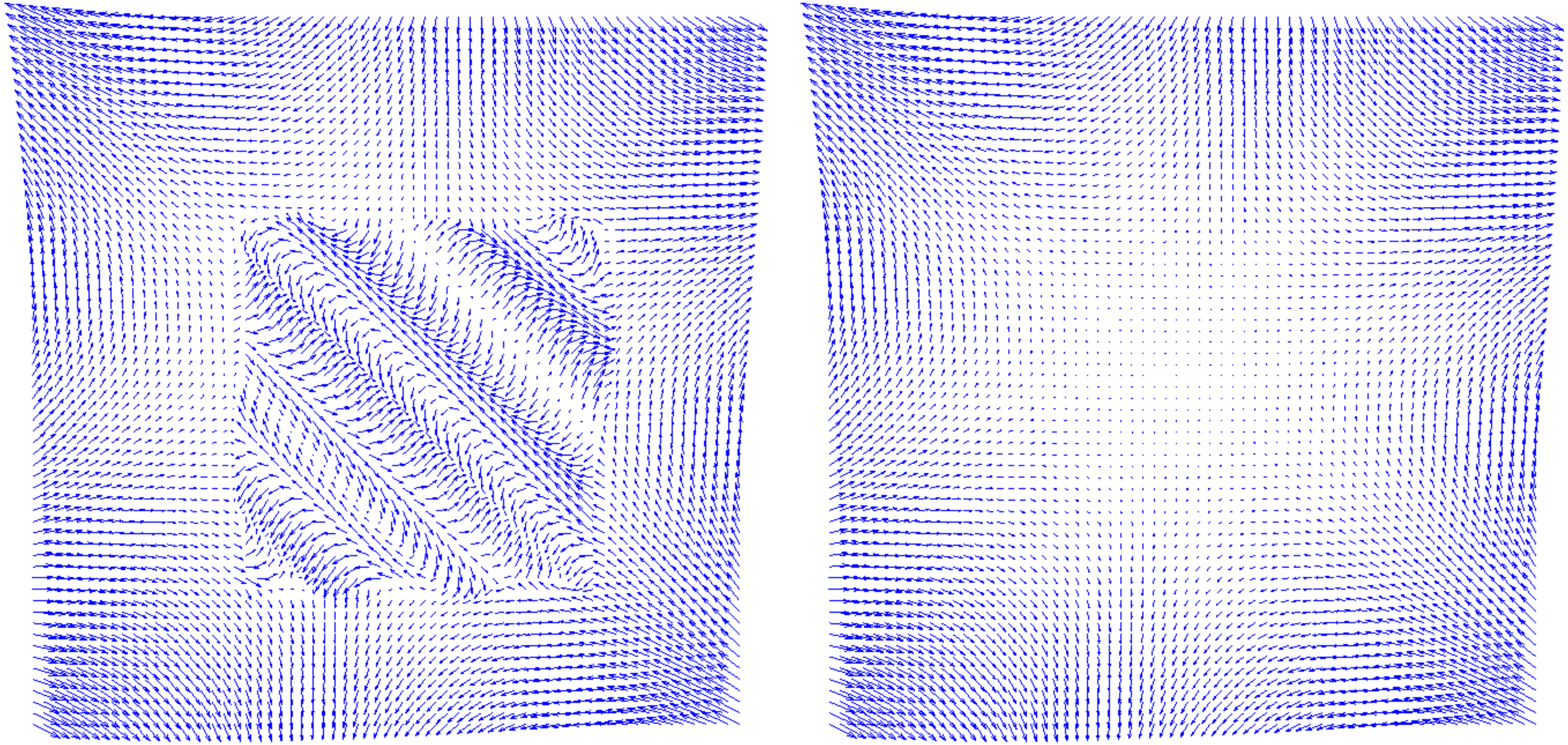


Standard TV-term, v-component

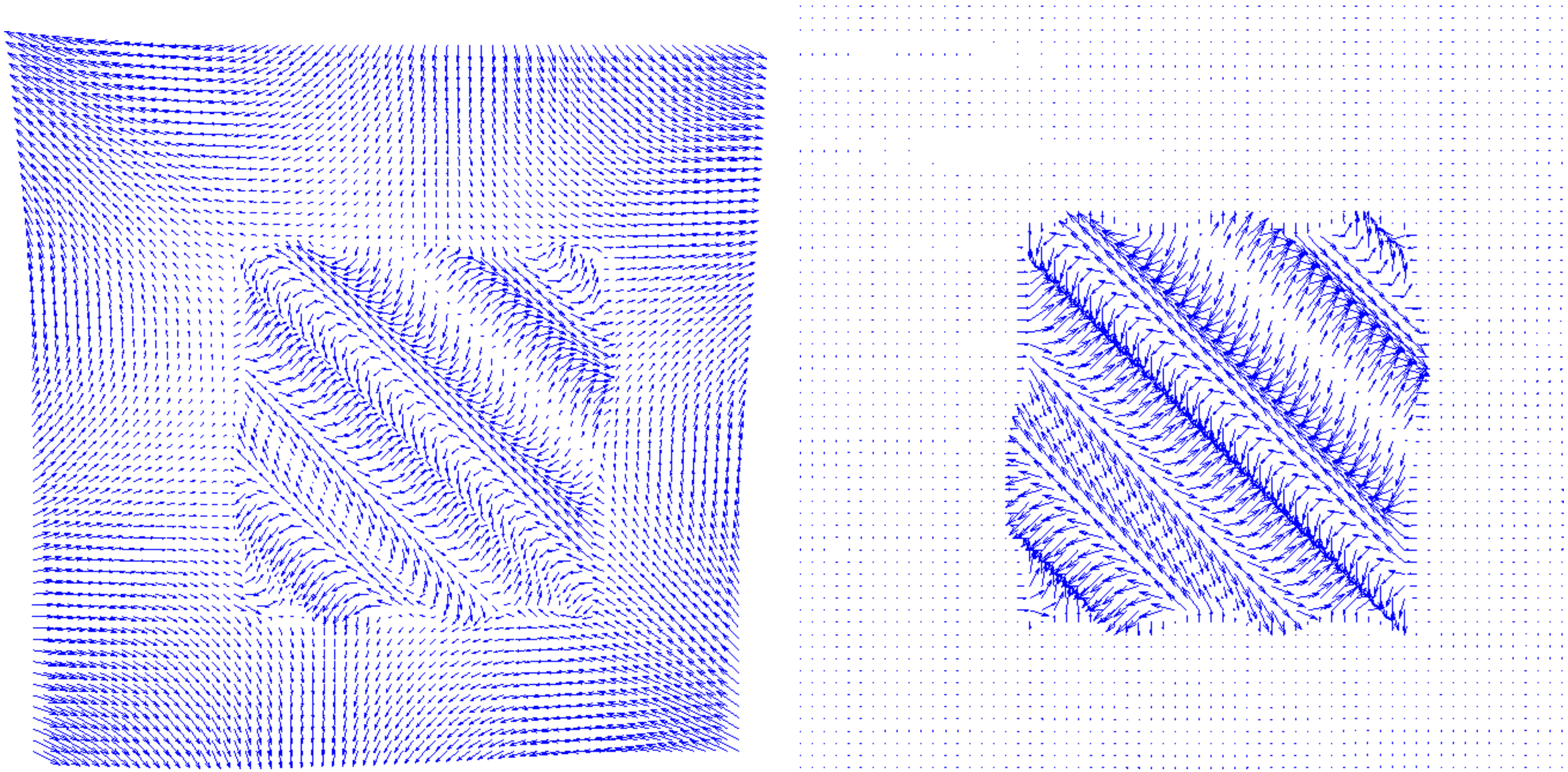




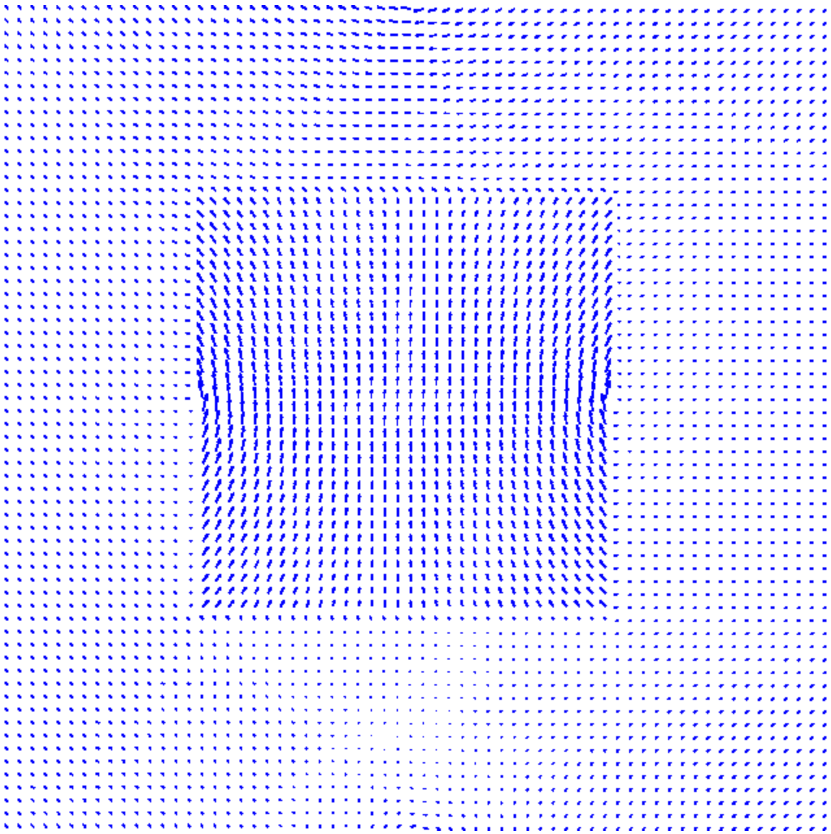
Alternative TV-term, u-component

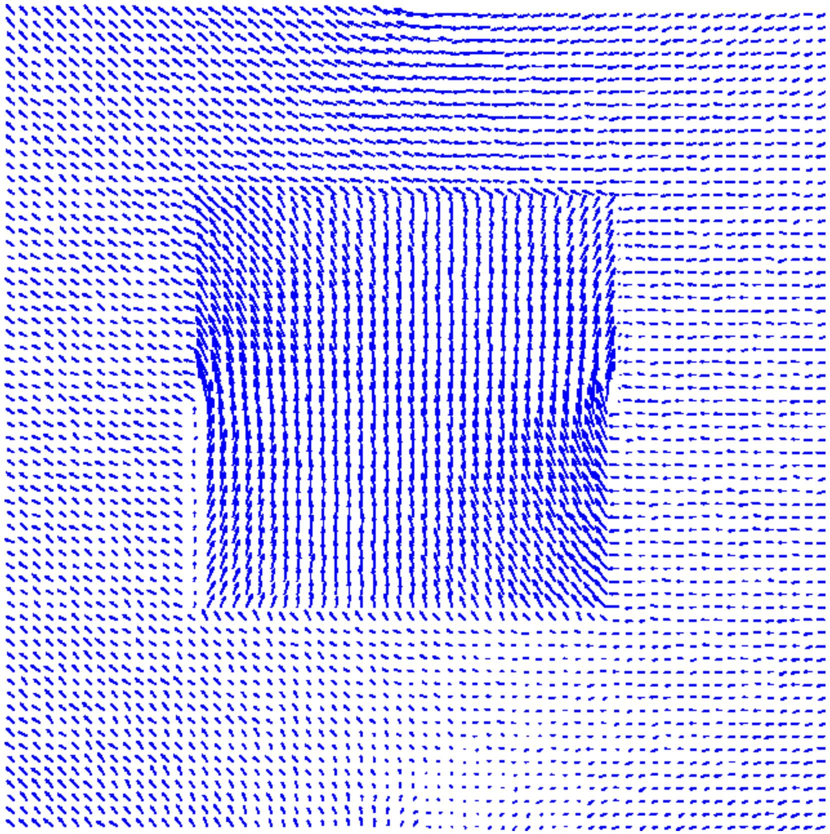


Alternative TV-term, v-component

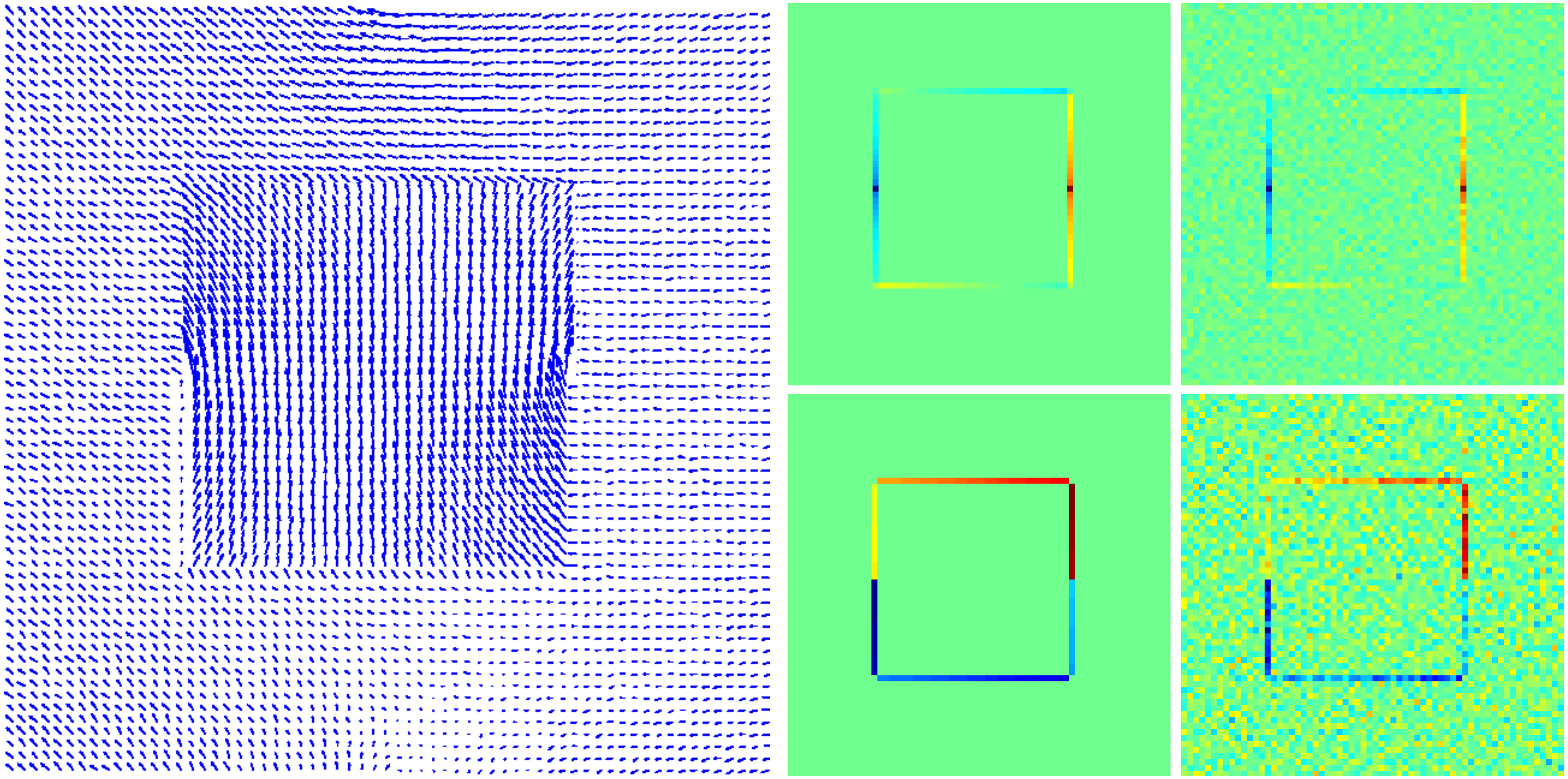


Denoising: Input data

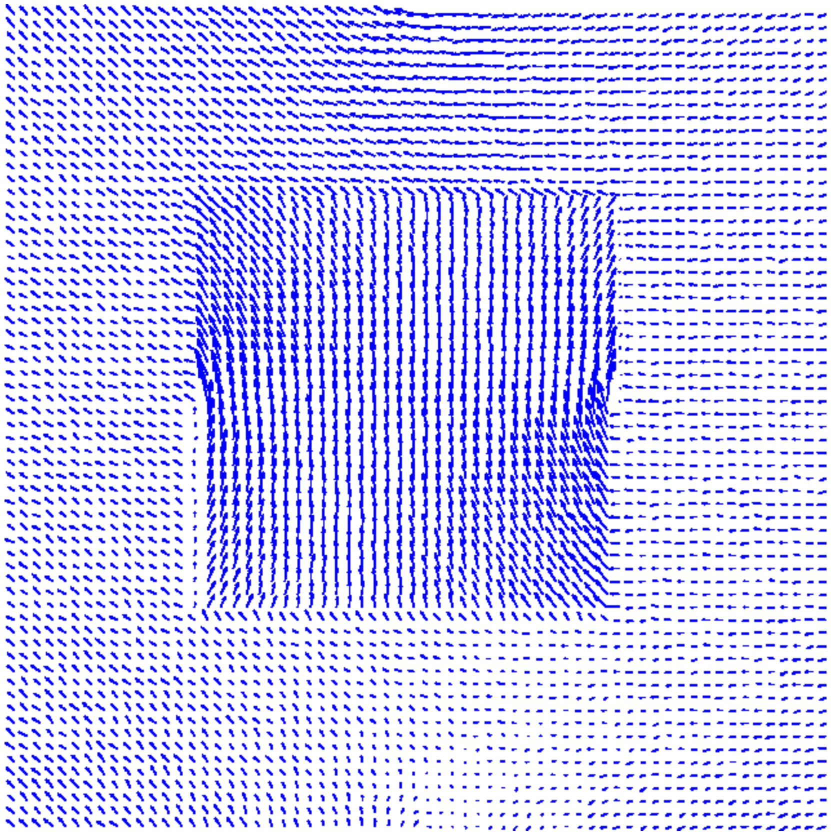


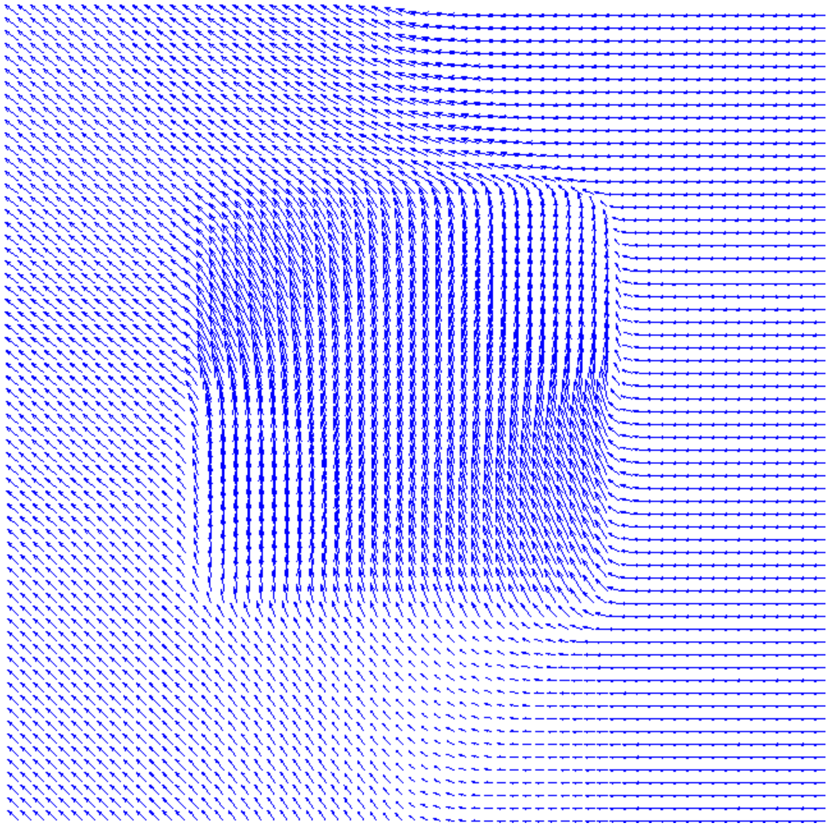


Denoising: Input data

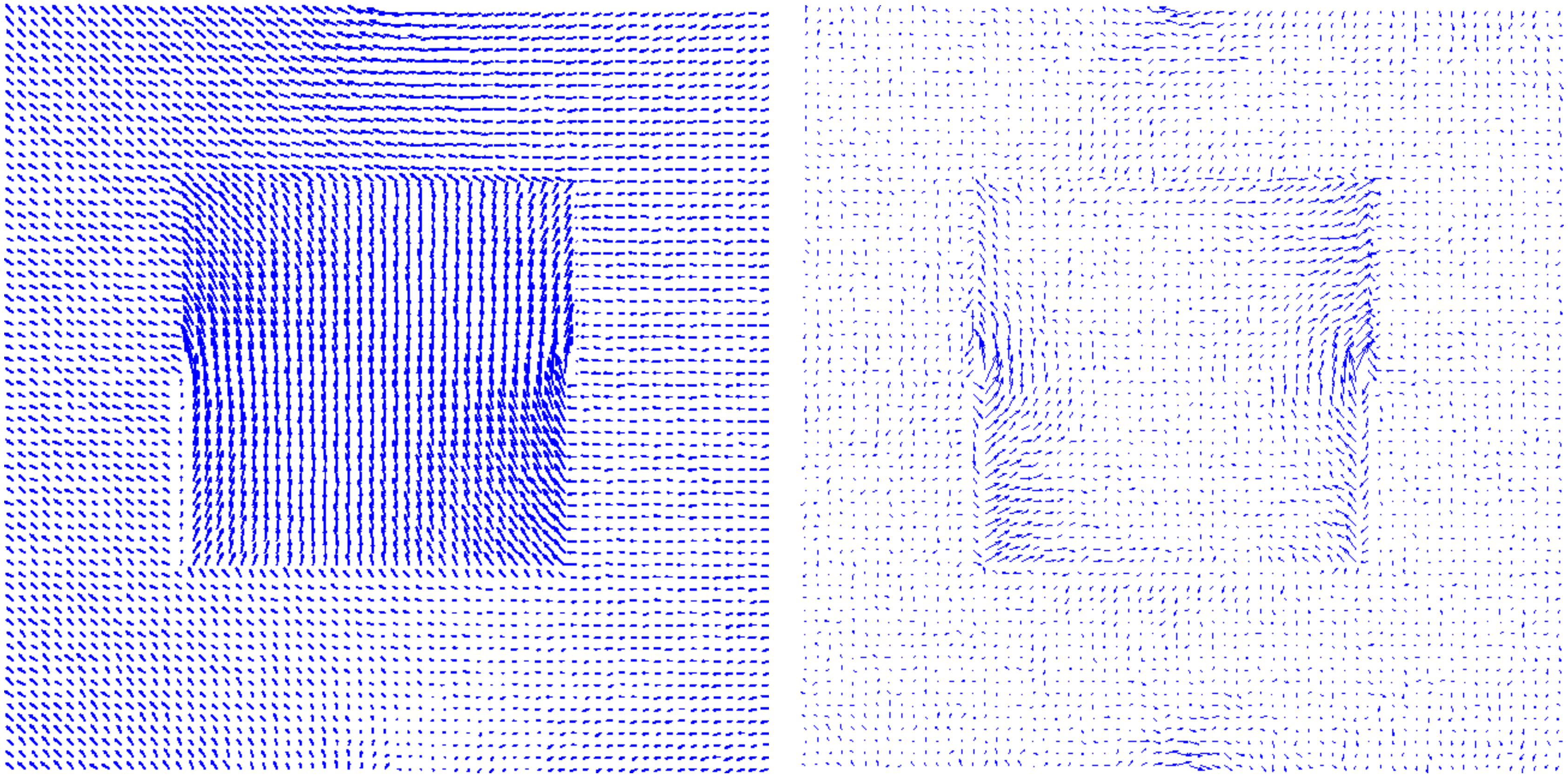


Standard TV-term, u-component

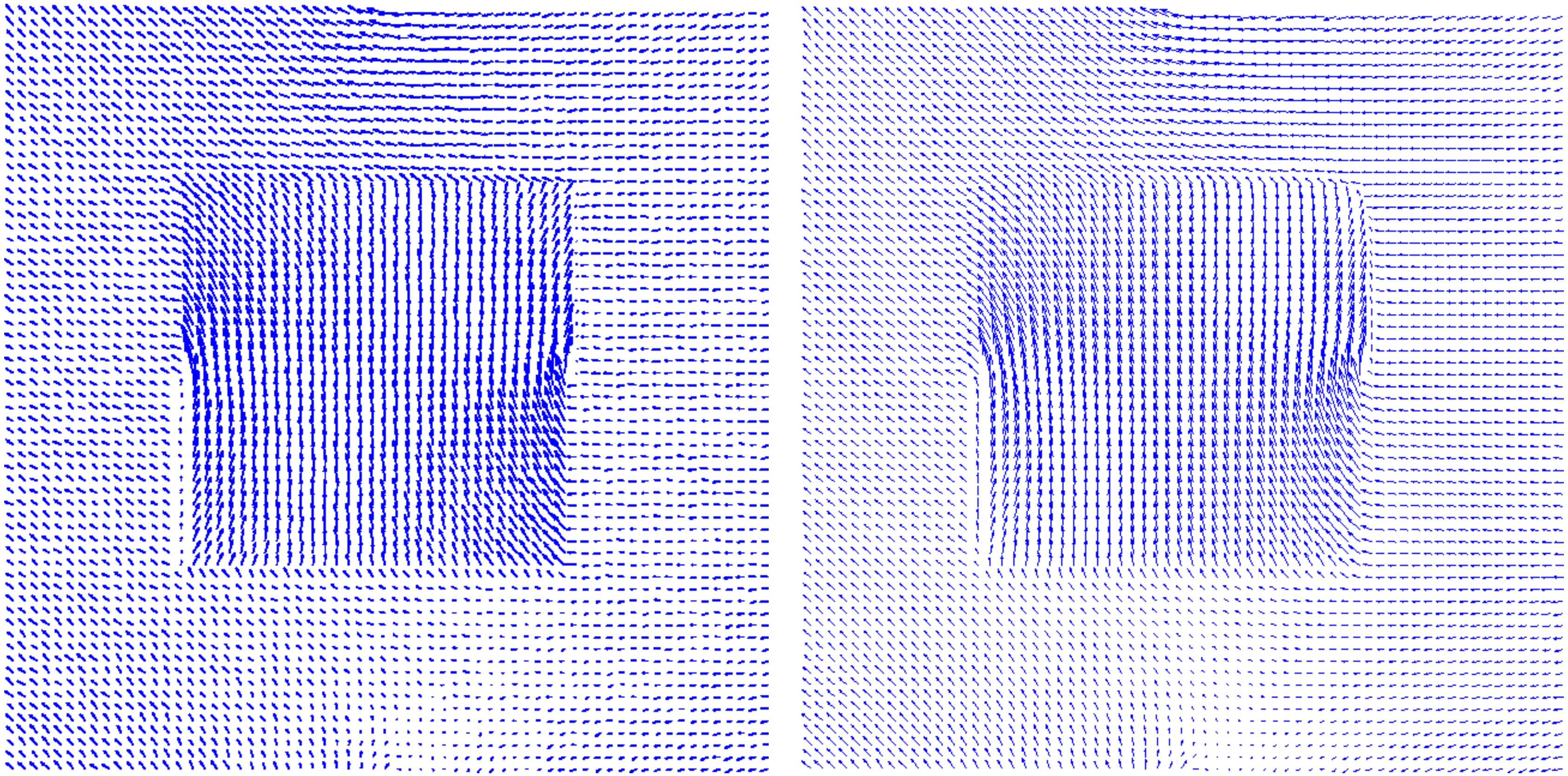




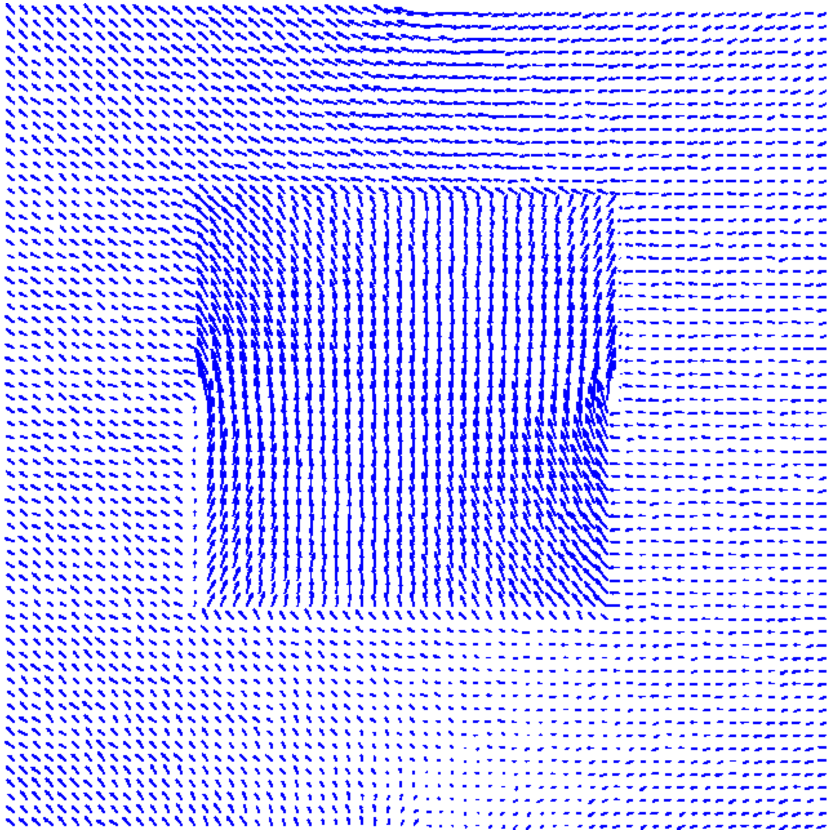
Standard TV-term, v-component

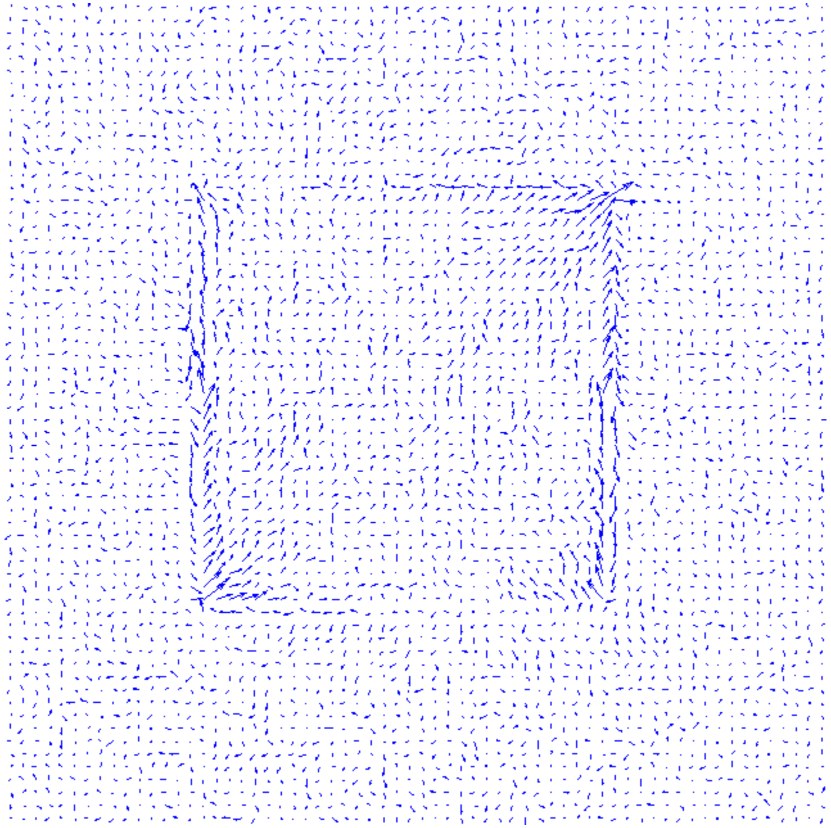


Alternative TV-term, u-component

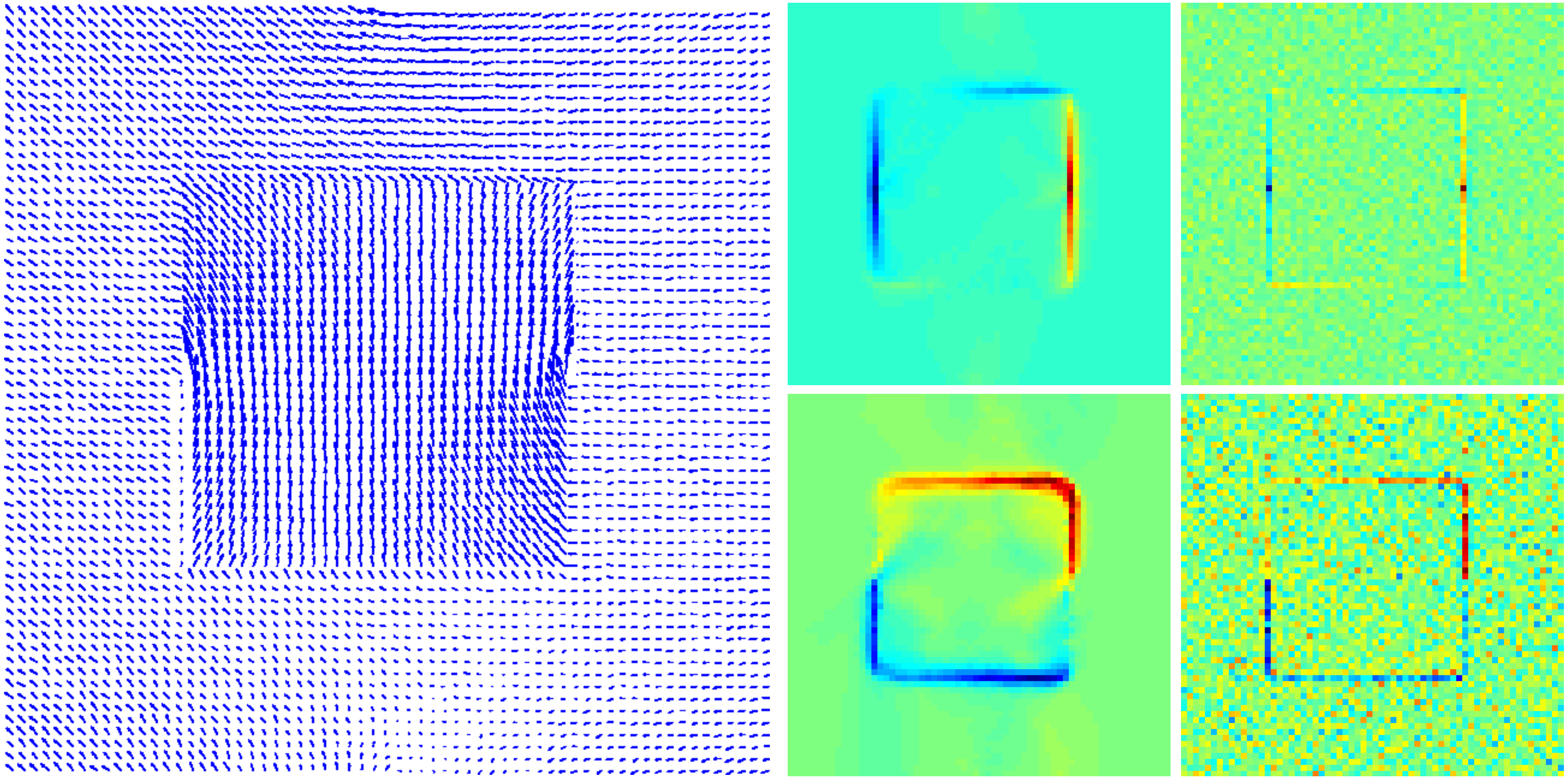


Alternative TV-term, v-component

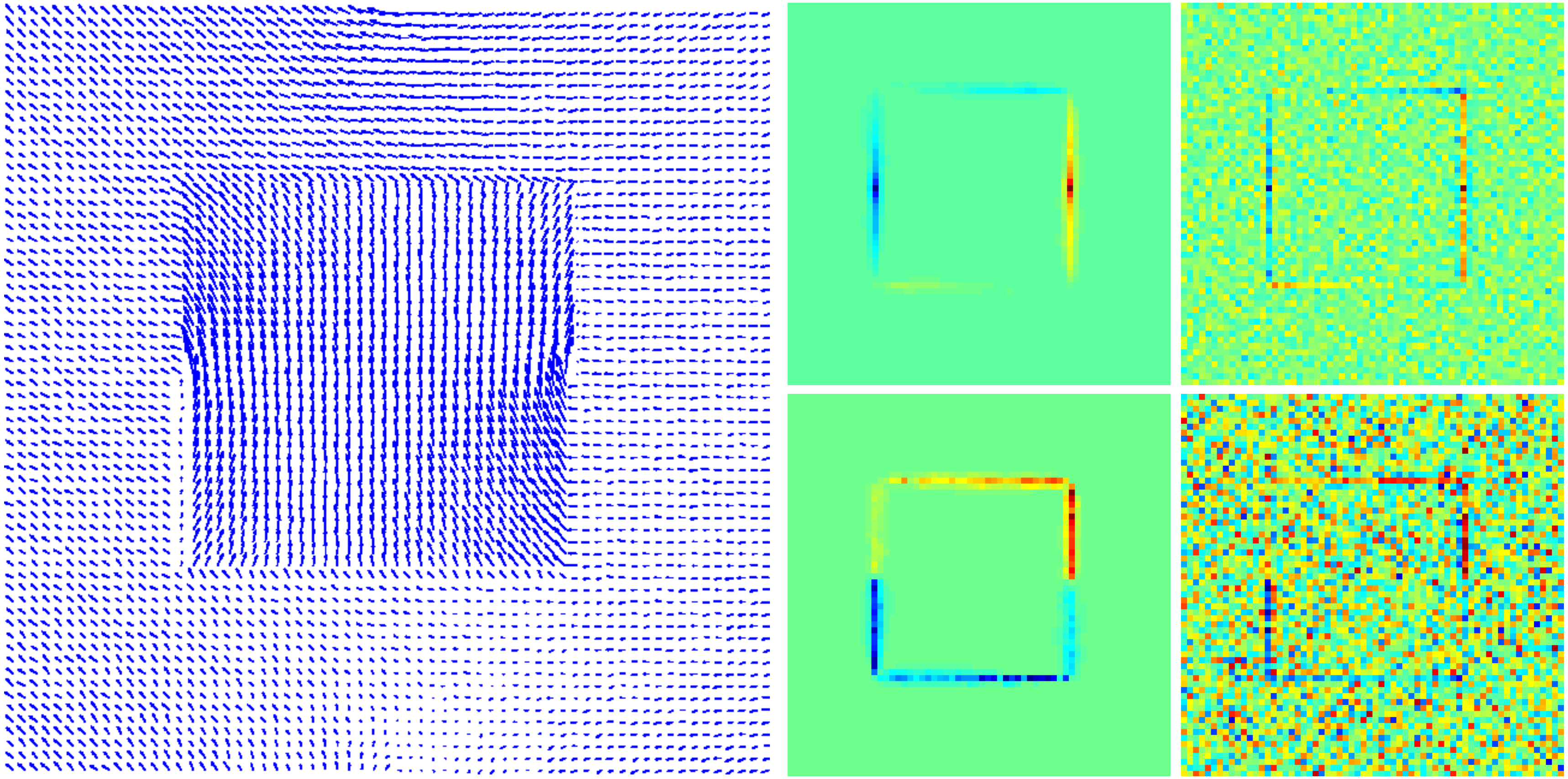




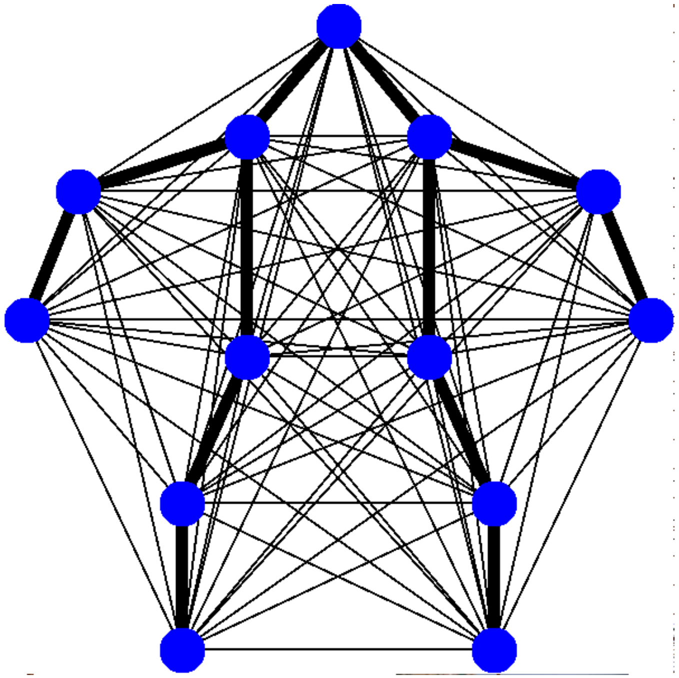
Standard TV-term

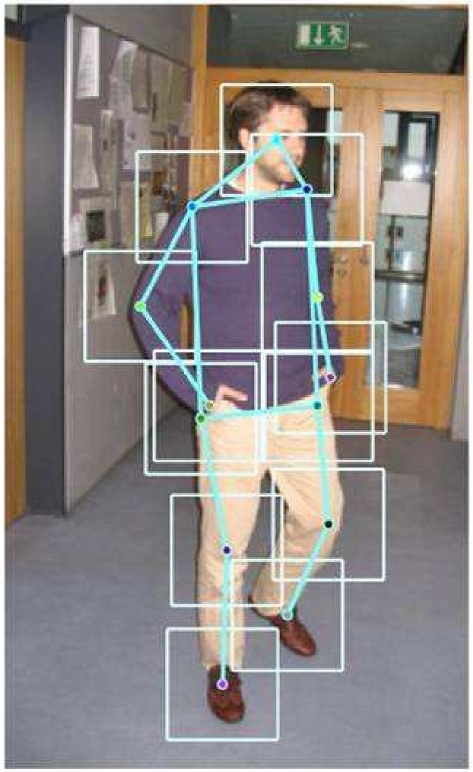


Alternative TV-term



MRF-Inference by DC-Programming





- $5 \dots 25$ vertices
- $20 \dots 100$ states

Fairly different from grid-graphs.

Related work: *Ravikumar&Lafferty*, *audience*.

MAP inference

Overcomplete representation (*Wainwright&Jordan*)

MAP-problem is equivalent to

$$\sup_{\mu \in \mathcal{M}_G} \langle \theta, \mu \rangle$$

Ravikumar&Lafferty:

$$\mu_{s;i} = \tau_{s;i} \ , \quad \forall s \in V \ ,$$

$$\mu_{st;ij} \approx \tau_{s;i} \tau_{t;j} \ , \quad \forall st \in E \ .$$

Reversing signs $\rightarrow$ “small” *non*-convex QP

$$\min_{\tau \in \Delta} \left\{ \frac{1}{2} \langle \tau, Q\tau \rangle + \langle q, \tau \rangle \right\}$$

$$\Delta = \left\{ \tau \ \Big| \ \sum_i \tau_{s;i} = 1 \, , \ \forall s \in V \right\}$$

Complexity

Method	# Constraints	# Variables
A^*	0	$\lvert V\rvert$
DC / QP	$\lvert V\rvert$	$\lvert V\rvert \cdot L$
SOCP	$\lvert V\rvert + \lvert V\rvert \cdot L + 3K$	$\lvert V\rvert \cdot L + 2 \cdot K$
LP	$\lvert V\rvert \cdot L + \lvert E\rvert \cdot L^2$	$\lvert V\rvert \cdot L + \lvert E\rvert \cdot L^2$
	$+\lvert V\rvert \cdot L + 2 \cdot \lvert E\rvert \cdot L$	
TRBP / BP	0	$2 \cdot \lvert E\rvert \cdot L$

DC-programming

Decomposition of the objective function

$$f(x) = g(x) - h(x) \ , \quad g, h \text{ proper, lsc, convex} \ .$$

Affine majorization of the concave part $\quad \rightarrow \quad$ two-step iteration

$$y^k \in \partial h(x^k) \ ,$$
$$x^{k+1} \in \partial g^*(y^k) \ .$$

Different decompositions $\quad \rightarrow \quad$ trade-off between

- complexity of each step,
- number of iterations.

Local convergence, empirically: “good” optima.

Immediate decompositions $Q = Q_1 - Q_2$

Spectral decomposition

$$Q_1 = V \operatorname{diag}\big(\ldots, \max\{0, \lambda_i(Q)\}, \ldots\big) V^\top$$
$$Q_2 = V \operatorname{diag}\big(\ldots, \max\{0, -\lambda_i(Q)\}, \ldots\big) V^\top$$

Lower eigenvalue bound $d_{\min} < \lambda_{\min}(Q) < 0$

$$Q_1 = Q - d_{\min} I$$
$$Q_2 = -d_{\min} I$$

Upper eigenvalue bound $0 < \lambda_{\max}(Q) < d_{\max}$

$$Q_1 = d_{\max} I$$
$$Q_2 = d_{\max} I - Q$$

Experiments

- $5 \dots 20$ vertices,
- $3 \dots 50$ labels,
- 100 random repetitions.

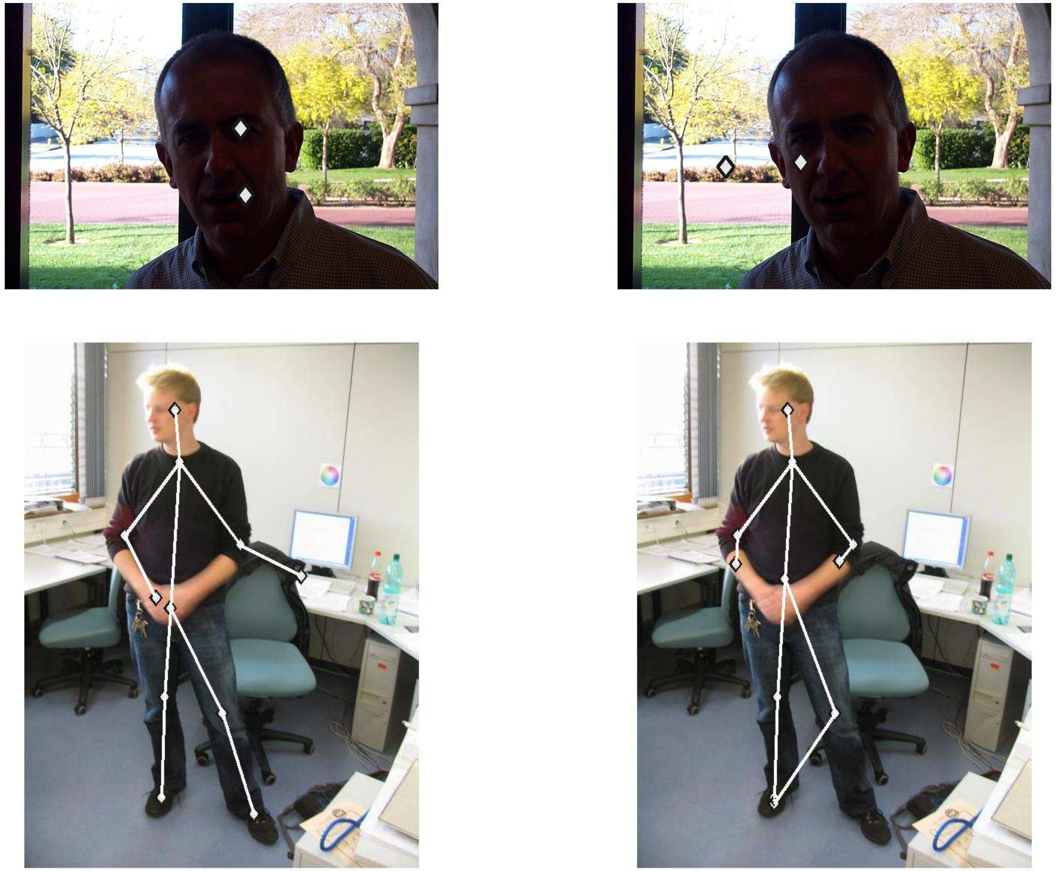
	Exp. A	Exp. B	Exp. C
θ_s	$-\log(\mathsf{U}(0,1))$	$-\log(1)$	$-\log(1)$
θ_{st}	$-\log(\mathsf{U}(0,1))$	$-\log(\mathsf{U}(0,1))$	$-\log(\{0.1, 1\})$

Preliminary conclusion

- DC performs best (next to A^*) for A,B-experiments.
  More runtime (MatLab/Mosek). C–tuning might change this.
- BP/TRBP slightly superior for C-experiments (issue of convergence).
- Real difficult problems (human): BR $\leq$ DC $\leq$ TRBP.
- Rectified convex QP performs worse.
- SOCP is expensive for larger problems (e.g., 10 nodes, 50 states).

Observation

DC-suboptimal solutions make often more sense than those computed with TRBP.



Decomposition bounds

Lower bound of the standard non-convex QP

$$l_Q = \min_{x \in \Delta} \{ \langle x, Qx \rangle \}$$

Optimizing the decomposition bound $Q = S + (Q - S)$ (*Anstreicher&Burer*)

$$\sup_S \{ l_S + l_{Q-S} \, , \; S \succeq 0 \, , \; S - Q \succeq 0 \}$$

$$= \min_X \{ \langle Q, X \rangle \, , \; \langle E, X \rangle = 1 \, , \; Xe \geq 0 \, , \; \mathrm{Diag}(Xe) \succeq X \succeq 0 \}$$

Is this the best possible bound? – No!

Alternative *dominating* bounds exist.

On TV Regularization of Image Flows and MRF-Inference by DC-Programming”

Jörg Kappes, Christoph Schnörr, Jing Yuan
University of Heidelberg

“Graph Cuts and Related Discrete
or Continuous Optimization Problems”
IPAM, Feb 25-29, 2008