Sublabel-Accurate Discretization of Nonconvex Free-Discontinuity Problems Supplementary Material

Thomas Möllenhoff Daniel Cremers
Technical University of Munich
{thomas.moellenhoff,cremers}@tum.de

Proposition 1. *For concave $\kappa : \mathbb{R}_0^+ \to \mathbb{R}$ with $\kappa(a) = 0 \Leftrightarrow a = 0$, the constraints*

$$\begin{aligned}&\Big\|(1-\alpha)\hat{\varphi}_x(i) + \sum_{l=i+1}^{j-1} \hat{\varphi}_x(l) + \beta\hat{\varphi}_x(j)\Big\| \\ &\quad\leq \frac{\kappa(\gamma_j^\beta - \gamma_i^\alpha)}{h}, \; \forall 1 \leq i \leq j \leq k, \alpha, \beta \in [0,1],\end{aligned} \tag{1}$$

are equivalent to

$$\Big\|\sum_{l=i}^{j} \hat{\varphi}_x(l)\Big\| \leq \frac{\kappa(\gamma_{j+1} - \gamma_i)}{h}, \forall 1 \leq i \leq j \leq k. \tag{2}$$

Proof. The implication (1) ⇒ (2) clearly holds. Let us now assume the constraints (2) are fulfilled. First we show that the constraints (1) also hold for $\alpha \in [0,1]$ and $\beta \in \{0,1\}$. First, we start with $\beta = 0$:

$$\begin{aligned}&\|(1-\alpha)\hat{\varphi}_x(i) + \sum_{l=i+1}^{j-1} \hat{\varphi}_x(l)\| = \\ &\|(1-\alpha)\sum_{l=i}^{j-1} \hat{\varphi}_x(l) + \alpha \sum_{l=i+1}^{j-1} \hat{\varphi}_x(l)\| \leq \\ &(1-\alpha)\|\sum_{l=i}^{j-1} \hat{\varphi}_x(l)\| + \alpha\|\sum_{l=i+1}^{j-1} \hat{\varphi}_x(l)\| \overset{\text{by (2)}}{\leq} \\ &(1-\alpha)\frac{1}{h}\kappa(\gamma_j - \gamma_i) + \alpha\frac{1}{h}\kappa(\gamma_j - \gamma_{i+1}) \overset{\text{concavity}}{\leq} \\ &\frac{1}{h}\left(\kappa((1-\alpha)(\gamma_j - \gamma_i) + \alpha(\gamma_j - \gamma_{i+1})\right) = \frac{1}{h}\kappa(\gamma_j^0 - \gamma_i^\alpha).\end{aligned} \tag{3}$$

In the same way, it can be shown that for $\beta = 1$ we have:

$$\|(1-\alpha)\hat{\varphi}_x(i) + \sum_{l=i+1}^{j-1} \hat{\varphi}_x(l) + 1 \cdot \hat{\varphi}_x(j)\| \leq \frac{1}{h}\kappa(\gamma_j^1 - \gamma_i^\alpha). \tag{4}$$

We have shown the constraints to hold for $\alpha \in [0,1]$ and $\beta \in \{0,1\}$. Finally we show they also hold for $\beta \in [0,1]$:

$$\begin{aligned}&\|(1-\alpha)\hat{\varphi}_x(i) + \sum_{l=i+1}^{j-1} \hat{\varphi}_x(l) + \beta\hat{\varphi}_x(j)\| = \\ &\|(1-\alpha)\hat{\varphi}_x(i) + (1-\beta)\sum_{l=i+1}^{j-1} \hat{\varphi}_x(l) + \beta\sum_{l=i+1}^{j} \hat{\varphi}_x(l)\| = \\ &\|(1-\beta)\left((1-\alpha)\hat{\varphi}_x(i) + \sum_{l=i+1}^{j-1} \hat{\varphi}_x(l)\right) + \\ &\beta\left((1-\alpha)\hat{\varphi}_x(i) + \sum_{l=i+1}^{j} \hat{\varphi}_x(l)\right)\| \leq \\ &(1-\beta)\|(1-\alpha)\hat{\varphi}_x(i) + \sum_{l=i+1}^{j-1} \hat{\varphi}_x(l)\| + \\ &\beta\|(1-\alpha)\hat{\varphi}_x(i) + \sum_{l=i+1}^{j} \hat{\varphi}_x(l)\| \overset{(3),(4)}{\leq} \\ &\frac{1}{h}(1-\beta)\kappa(\gamma_j^0 - \gamma_i^\alpha) + \beta\kappa(\gamma_j^1 - \gamma_i^\alpha) \overset{\text{concavity}}{\leq} \\ &\frac{1}{h}\kappa((1-\beta)(\gamma_j^0 - \gamma_i^\alpha) + \beta(\gamma_j^1 - \gamma_i^\alpha)) = \frac{1}{h}\kappa(\gamma_j^\beta - \gamma_i^\alpha)\end{aligned} \tag{5}$$

Noticing that (2) is precisely (1) for $\alpha, \beta \in \{0,1\}$ (as $\kappa(a) = 0 \Leftrightarrow a = 0$) completes the proof. □

Proposition 2. *For convex one-homogeneous η the discretization with piecewise constant φ_t and φ_x leads to the traditional discretization as proposed in [2], except with min-pooled instead of sampled unaries.*

Proof. The constraints in [2, Eq. 18] have the form

$$\hat{\varphi}_t(i) \geq \eta^*(\hat{\varphi}_x(i)) - \rho(\gamma_i), \tag{6}$$

$$\Big\|\sum_{l=i}^{j} \hat{\varphi}_x(l)\Big\| \leq \kappa(\gamma_{j+1} - \gamma_i), \tag{7}$$

with $\rho(u) = \lambda(u-f)^2$, $\eta(g) = \|g\|^2$ and $\kappa(a) = \nu[\![a > 0]\!]$. The constraints (7) are equivalent to (2) up to a rescaling of

$\hat{\varphi}_x$ with h. For the constraints (6) (cf. [2, Eq. 18]), the unaries are sampled at the labels γ_i. The discretization with piecewise constant duals leads to a similar form, except for a min-pooling on dual intervals, $\forall 1 \leq i \leq k$:

$$\begin{aligned} \hat{\varphi}_t(i) &\geq \eta^*(\hat{\varphi}_x(i)) - \inf_{t \in [\gamma_i, \gamma_i^*]} \rho(t), \\ \hat{\varphi}_t(i+1) &\geq \eta^*(\hat{\varphi}_x(i)) - \inf_{t \in [\gamma_i^*, \gamma_{i+1}]} \rho(t). \end{aligned} \tag{8}$$

The similarity between (8) and (6) becomes more evident by assuming convex one-homogeneous η. Then (8) reduces to the following:

$$\begin{aligned} -\hat{\varphi}_t(1) &\leq \inf_{t \in [\gamma_1, \gamma_1^*]} \rho(t), \\ -\hat{\varphi}_t(i) &\leq \inf_{t \in \Gamma_i^*} \rho(t), \; \forall i \in \{2, \dots, \ell - 1\}, \\ -\hat{\varphi}_t(\ell) &\leq \inf_{t \in [\gamma_{\ell-1}^*, \gamma_\ell]} \rho(t), \end{aligned} \tag{9}$$

as well as

$$\hat{\varphi}_x(i) \in \mathsf{dom}(\eta^*), \forall i \in \{1, \dots, k\}. \tag{10}$$

□

Proposition 3. *The constraints*

$$\begin{aligned} \inf_{t \in \Gamma_i} \; & \hat{\varphi}_t(i) \frac{\gamma_{i+1} - t}{h} + \hat{\varphi}_t(i+1) \frac{t - \gamma_i}{h} \\ & + \rho(t) \geq \eta^*(\hat{\varphi}_x(i)). \end{aligned} \tag{11}$$

can be equivalently reformulated by introducing additional variables $a \in \mathbb{R}^k$, $b \in \mathbb{R}^k$, where $\forall i \in \{1, \dots, k\}$:

$$\begin{aligned} & r(i) = (\hat{\varphi}_t(i) - \hat{\varphi}_t(i+1))/h, \\ & a(i) + b(i) - (\hat{\varphi}_t(i)\gamma_{i+1} - \hat{\varphi}_t(x, i+1)\gamma_i)/h = 0, \\ & r(i) \geq \rho_i^*\left(a(i)\right), \hat{\varphi}_x(i) \geq \eta^*\left(b(i)\right), \end{aligned} \tag{12}$$

with $\rho_i(x,t) = \rho(x,t) + \delta\{t \in \Gamma_i\}$.

Proof. Rewriting the infimum in (11) as minus the convex conjugate of ρ_i, and multiplying the inequality with -1 the constraints become:

$$\begin{aligned} & \rho_i^*(r(i)) + \eta^*(\hat{\varphi}_x(i)) - \frac{\hat{\varphi}_t(i)\gamma_{i+1} - \hat{\varphi}_t(i+1)\gamma_i}{h} \leq 0, \\ & r(i) = (\hat{\varphi}(i) - \hat{\varphi}(i+1))/h. \end{aligned} \tag{13}$$

To show that (13) and (12) are equivalent, we prove that they imply each other. Assume (13) holds. Then without loss of generality set $a(i) = \rho_i^*(r(i)) + \xi_1$, $b(i) = \eta_i^*(\varphi_x(i)) + \xi_2$ for some $\xi_1, \xi_2 \geq 0$. Clearly, this choice fulfills (13). Since for $\xi_1 = \xi_2 = 0$ we have by assumption that

$$a(i) + b(i) - (\hat{\varphi}_t(i)\gamma_{i+1} - \hat{\varphi}_t(x, i+1)\gamma_i)/h \leq 0, \tag{14}$$

there exists some $\xi_1, \xi_2 \geq 0$ such that (12) holds.

Now conversely assume (12) holds. Since $a(i) \geq \rho_i^*\left(r(i)\right)$, $b(i) \geq \eta^*\left(\hat{\varphi}_x(i)\right)$, and

$$a(i) + b(i) - (\hat{\varphi}_t(i)\gamma_{i+1} - \hat{\varphi}_t(x, i+1)\gamma_i)/h = 0, \tag{15}$$

this directly implies

$$\rho_i^*(r(i)) + \eta^*(\hat{\varphi}_x(i)) - \frac{\hat{\varphi}_t(i)\gamma_{i+1} - \hat{\varphi}_t(i+1)\gamma_i}{h} \leq 0, \tag{16}$$

since the left-hand side becomes smaller by plugging in the lower bound. □

Proposition 4. *The discretization with piecewise linear φ_t and piecewise constant φ_x together with the choice $\eta(g) = \|g\|$ and $\kappa(a) = a$ is equivalent to the relaxation [1].*

Proof. Since $\eta(g) = \|g\|$, the constraints (11) become

$$\begin{aligned} & \inf_{t \in \Gamma_i} \; \hat{\varphi}_t(i) \frac{\gamma_{i+1} - t}{h} + \hat{\varphi}_t(i+1) \frac{t - \gamma_i}{h} + \rho(t) \geq 0. \\ & \varphi_x \in \mathsf{dom}(\eta^*). \end{aligned} \tag{17}$$

This decouples the constraints into data term and regularizer. The data term constraints can be written using the convex conjugate of $\rho_i = \rho + \delta\{\cdot \in \Gamma_i\}$ as the following:

$$\frac{\hat{\varphi}_t(i)\gamma_{i+1} - \hat{\varphi}_t(i+1)\gamma_i}{h} - \rho_i^*\left(\frac{\hat{\varphi}_t(i) - \hat{\varphi}_t(i+1)}{h}\right) \geq 0. \tag{18}$$

Let $\boldsymbol{v}_i = \hat{\varphi}_t(i) - \hat{\varphi}_t(i+1)$ and $q = \hat{\varphi}_t(1)$. Then we can write (18) as a telescope sum over the $\boldsymbol{v}_i$

$$q - \sum_{j=1}^{i-1} \boldsymbol{v}_j + \frac{\gamma_i}{h}\boldsymbol{v}_i - \rho_i^*\left(\frac{\boldsymbol{v}_i}{h}\right) \geq 0, \tag{19}$$

which is the same as the constraints in [1, Eq. 9,Eq. 22]. The cost function is given as

$$-\hat{\varphi}_t(1) - \sum_{i=1}^{k} \hat{v}(i)\left[\hat{\varphi}_t(i+1) - \hat{\varphi}_t(i)\right] = \langle \hat{v}, \boldsymbol{v} \rangle - q, \tag{20}$$

which is exactly the first part of [1, Eq. 21]. Finally, for the regularizer we get

$$\Big\|\sum_{l=i}^{j} \hat{\varphi}_x(l)\Big\| \leq \frac{|\gamma_{j+1} - \gamma_i|}{h}, \; \|\hat{\varphi}_x(i)\| \leq 1, \tag{21}$$

which clearly reduces to the same set as in [1, Proposition 5], by applying that proposition (and with the rescaling/substitution $p = h \cdot \varphi_x$).

□

Proposition 5. *The data term from [1] (which is in turn a special case of the discretization with piecewise linear φ_t) can be (pointwise) brought into the primal form*

$$\mathcal{D}(\widehat{v}) = \inf_{\substack{x_i \geq 0, \sum_i x_i = 1 \\ \widehat{v} = y/h + I^\top x}} \sum_{i=1}^{k} x_i \rho_i^{**}\left(\frac{y_i}{x_i}\right), \tag{22}$$

where $I \in \mathbb{R}^{k \times k}$ is a discretized integration operator.

Proof. The equivalence of the sublabel accurate data term proposed in [1] to the discretization with piecewise linear φ_t is established in Proposition 4 (cf. (19) and (20)). It is given pointwise as

$$\mathcal{D}(\widehat{v}) = \max_{\boldsymbol{v}, q} \langle \boldsymbol{v}, \widehat{v} \rangle - q - \sum_{i=1}^{k} \delta\left\{ \left(\frac{\boldsymbol{v}_i}{h}, [q\mathbf{1}_k - I\boldsymbol{v}]_i\right) \in \mathsf{epi}(\rho_i^*) \right\}, \tag{23}$$

where $\widehat{v} \in \mathbb{R}^k, \boldsymbol{v} \in \mathbb{R}^k, q \in \mathbb{R}$, and k is the number of pieces and $\mathbf{1}_k \in \mathbb{R}^k$ is the vector consisting only of ones. Furthermore, $\rho_i(t) = \rho(t) + \delta\{t \in \Gamma_i\}, \mathsf{dom}(\rho_i) = \Gamma_i = [\gamma_i, \gamma_{i+1}]$. The integration operator $I \in \mathbb{R}^{k \times k}$ is defined as

$$I = \begin{bmatrix} -\frac{\gamma_1}{h} & & & \\ 1 & -\frac{\gamma_2}{h} & & \\ & & \ddots & \\ 1 & \dots & 1 & -\frac{\gamma_k}{h} \end{bmatrix}. \tag{24}$$

Using convex duality, and the substitution $h\tilde{v} = \boldsymbol{v}$ we can rewrite (23) as

$$\min_{x} \max_{\tilde{v}, q, z} \langle \tilde{v}, h \cdot \widehat{v} \rangle - q - \langle x, z - (q\mathbf{1}_k - hI\tilde{v}) \rangle - \sum_{i=1}^{k} \delta\left\{ (\tilde{v}_i, z_i) \in \mathsf{epi}(\rho_i^*) \right\}, \tag{25}$$

The convex conjugate of $F_i(z, v) = \delta\{(v, -z) \in \mathsf{epi}(\rho_i^*)\}$ is the lower-semicontinuous envelope of the perspective [3, Section 15], and since $\rho_i : \Gamma_i \to \mathbb{R}$ has bounded domain, is given as the following (cf. also [5, Appendix 3])

$$F_i^*(x, y) = \begin{cases} x\rho_i^{**}(y/x), & \text{if } x > 0, \\ 0, & \text{if } x = 0 \wedge y = 0, \\ \infty, & \text{if } x < 0 \vee (x = 0 \wedge y \neq 0). \end{cases} \tag{26}$$

Thus with the convention that $0/0 = 0$ equation (25) can be rewritten as convex conjugates:

$$\min_{x} \left(\max_{q} q(\mathbf{1}_k^\top x) - q \right) + \left(\max_{\tilde{v}, z} \langle \tilde{v}, h \cdot (\widehat{v} - I^\top x) \rangle + \langle -z, x \rangle - \sum_{i=1}^{k} F_i(-z_i, \tilde{v}_i) \right) = \min_{x} \delta\left\{ \sum_i x_i = 1 \right\} + \sum_i F_i^*\left(x_i, \left[h(\widehat{v} - I^\top x)\right]_i\right). \tag{27}$$

Hence we have that

$$\mathcal{D}(\widehat{v}) = \min_{\substack{x, y \\ y = h(\widehat{v} - I^\top x) \\ x_i \geq 0 \\ \sum_i x_i = 1 \\ y_i / x_i \in \mathsf{dom}(\rho_i^{**})}} \sum_i x_i \rho_i^{**}\left(\frac{y_i}{x_i}\right), \tag{28}$$

which can be rewritten in the form (23). □

Proposition 6. *Let $\gamma = \kappa(\gamma_2 - \gamma_1)$ and $\ell = 2$. The approximation with piecewise linear φ_t and piecewise constant φ_x of the continuous optimization problem*

$$\inf_{v \in \mathcal{C}} \sup_{\varphi \in \mathcal{K}} \int_{\Omega \times \mathbb{R}} \langle \varphi, Dv \rangle. \tag{29}$$

is equivalent to

$$\inf_{u : \Omega \to \Gamma} \int_{\Omega} \rho^{**}(x, u(x)) + (\eta^{**} \,\square\, \gamma \| \cdot \|)(\nabla u(x)) \,\mathrm{d}x, \tag{30}$$

where $(\eta \,\square\, \gamma\| \cdot \|)(x) = \inf_y \; \eta(x - y) + \gamma\|y\|$ denotes the infimal convolution (cf. [3, Section 5]).

Proof. Plugging in the representations for piecewise linear φ_t and piecewise constant φ_x we have the coefficient functions $\hat{v} : \Omega \to [0, 1]$, $\hat{\varphi}_t : \Omega \times \{1, 2\} \to \mathbb{R}$, $\hat{\varphi}_x : \Omega \to \mathbb{R}^n$ and the following optimization problem:

$$\begin{aligned} \inf_{\hat{v}} \sup_{\hat{\varphi}_x, \hat{\varphi}_t} \int_{\Omega} & - \hat{\varphi}_t(x, 1) - \hat{v}(x) \left[\hat{\varphi}_t(x, 2) - \hat{\varphi}_t(x, 1)\right] \\ & - h \cdot \hat{v}(x) \cdot \mathrm{Div}_x \, \hat{\varphi}_x(x) \,\mathrm{d}x \\ & \text{subject to} \\ \inf_{t \in \Gamma} \hat{\varphi}_t(x, 1) & \frac{\gamma_2 - t}{h} + \hat{\varphi}_t(x, 2) \frac{t - \gamma_1}{h} + \rho(x, t) \geq \eta^*(x, \hat{\varphi}_x(x)) \\ \|\hat{\varphi}_x(x)\| & \leq \kappa(\gamma_2 - \gamma_1) =: \gamma. \end{aligned} \tag{31}$$

Using the convex conjugate of $\rho : \Omega \times \Gamma \to \mathbb{R}$ (in its second argument), we rewrite the first constraint as

$$\frac{\hat{\varphi}_t(x, 1)\gamma_2 - \hat{\varphi}_t(x, 2)\gamma_1}{h} \geq \rho^*\left(x, \frac{\hat{\varphi}_t(x, 1) - \hat{\varphi}_t(x, 2)}{h}\right) + \eta^*(x, \hat{\varphi}_x(x)). \tag{32}$$

Using the substitution $\tilde{\varphi}(x) = \frac{\hat{\varphi}_t(x,1) - \hat{\varphi}_t(x,2)}{h}$ we can reformulate the constraints as

$$\hat{\varphi}_t(x,1) \geq \rho^*(x, \tilde{\varphi}(x)) + \eta^*(x, \hat{\varphi}_x(x)) - \gamma_1 \tilde{\varphi}(x), \quad (33)$$

and the cost function as

$$\sup_{\tilde{\varphi}, \hat{\varphi}_t, \hat{\varphi}_x} \int_\Omega -\hat{\varphi}_t(x,1) + h\hat{v}(x)\tilde{\varphi}(x) - h\hat{v}(x) \operatorname{Div}_x \hat{\varphi}_x(x) \mathrm{d}x. \quad (34)$$

The pointwise supremum over $-\hat{\varphi}_t(x,1)$ is attained where the constraint (33) is sharp, which means we can pull it into the cost function to arrive at

$$\sup_{\tilde{\varphi}, \hat{\varphi}_x} \int_\Omega -\rho^*(x, \tilde{\varphi}(x)) - \eta^*(x, \hat{\varphi}_x(x)) - \delta\{\|\hat{\varphi}_x(x) \leq \gamma\|\} + \gamma_1 \tilde{\varphi}(x) + h\hat{v}(x)\tilde{\varphi}(x) - h\hat{v}(x) \operatorname{Div}_x \hat{\varphi}_x(x) \mathrm{d}x, \quad (35)$$

where we wrote the second constraint in (31) as an indicator function. As the supremum decouples in $\tilde{\varphi}$ and $\hat{\varphi}_x$, we can rewrite it using convex (bi-)conjugates, by interchanging integration and supremum (cf. [4, Theorem 14.60]):

$$\begin{aligned} \sup_{\tilde{\varphi}} \int_\Omega \gamma_1 \tilde{\varphi}(x) + h\hat{v}(x)\tilde{\varphi}(x) - \rho^*(x, \tilde{\varphi}(x)) \mathrm{d}x &= \\ \int_\Omega \sup_{\tilde{\varphi}} \ \gamma_1 \tilde{\varphi} + h\hat{v}(x)\tilde{\varphi} - \rho^*(x, \tilde{\varphi}) \mathrm{d}x &= \\ \int_\Omega \rho^{**}(x, \gamma_1 + h\hat{v}(x)) \, \mathrm{d}x. & \end{aligned} \quad (36)$$

For the part in $\hat{\varphi}_x$ we assume that $\hat{v}$ is sufficiently smooth and apply partial integration ($\hat{\varphi}_x$ vanishes on the boundary), and then perform a similar calculation to the previous one:

$$\begin{aligned} &\sup_{\hat{\varphi}_x} \int_\Omega -(\eta^* + \delta\{\|\cdot\| \leq \gamma\})(x, \hat{\varphi}_x(x)) - \\ &\qquad h\hat{v}(x) \operatorname{Div}_x \hat{\varphi}_x(x) \mathrm{d}x = \\ &\sup_{\hat{\varphi}_x} \int_\Omega -(\eta^* + \delta\{\|\cdot\| \leq \gamma\})(x, \hat{\varphi}_x(x)) + \\ &\qquad h\langle \nabla_x \hat{v}(x), \hat{\varphi}_x(x) \rangle \mathrm{d}x = \\ &\int_\Omega \sup_{\hat{\varphi}_x} -(\eta^* + \delta\{\|\cdot\| \leq \gamma\})(x, \hat{\varphi}_x) + \\ &\qquad h\langle \nabla_x \hat{v}(x), \hat{\varphi}_x \rangle \mathrm{d}x = \\ &\int_\Omega (\eta^* + \delta\{\|\cdot\| \leq \gamma\})^*(x, h\nabla_x \hat{v}(x)) \mathrm{d}x = \\ &\int_\Omega (\eta^{**} \,\square\, \gamma\|\cdot\|)(x, h\nabla_x \hat{v}(x)) \mathrm{d}x = \\ &\int_\Omega (\eta \,\square\, \gamma\|\cdot\|)(x, h\nabla_x \hat{v}(x)) \mathrm{d}x. \end{aligned} \quad (37)$$

Here we used also the result that $(f^* + g)^* = f^{**} \,\square\, g^*$ [4, Theorem 11.23]. Combining (36) and (37) and using the substitution $u = \gamma_1 + h\hat{v}$, we finally arrive at:

$$\int_\Omega \rho^{**}(x, u(x)) + (\eta^{**} \,\square\, \gamma\|\cdot\|)(x, \nabla u(x)) \, \mathrm{d}x, \quad (38)$$

which is the same as (30). □

References

[1] T. Möllenhoff, E. Laude, M. Moeller, J. Lellmann, and D. Cremers. Sublabel-accurate relaxation of nonconvex energies. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 2016. 2, 3

[2] T. Pock, D. Cremers, H. Bischof, and A. Chambolle. An algorithm for minimizing the piecewise smooth Mumford-Shah functional. In *Proceedings of the IEEE International Conference on Computer Vision, ICCV*, 2009. 1, 2

[3] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, 1996. 3

[4] R. T. Rockafellar, R. J.-B. Wets, and M. Wets. *Variational analysis*. Springer, 1998. 4

[5] C. Zach and P. Kohli. A convex discrete-continuous approach for Markov random fields. In *Proceedings of the European Conference on Computer Vision, ECCV*, 2014. 3