Computing the Critical Temperature for the Dense Gaussian CRFs

In this document, we first compute analytically the phase transition temperature parameter T_c of 6.2 where the KL-Divergence stops being convex. In the first part *Analytical Derivation*, we make strong assumptions in order to be able to obtain a closed form estimation of T_c. We then explain how this result helps understanding real cases. In the second part *Experimental Analysis*, in order to justify our assumptions, we run experiments under three regimes, one where our assumptions are strictly verified, one which corresponds to a real-life scenario and an intermediate one. This set of experiments shows that our strong assumptions provide a valuable insight for practical applications.

1 Analytical derivation

Let us take probability distribution P to be defined by a dense Gaussian CRF [2]. In order to make computation tractable, we assume that the RGB distance between pixels is uniform and equal to d_{rgb}. Therefore the RGB Kernel is constant with value

$$\theta_{rgb} = e^{\frac{-d_{rgb}^2}{2\sigma_{rgb}}} . \tag{1}$$

We consider the case where we have only two possible labels and the same unary potential on all the variables. Even if this assumption sounds strong, we can expect them to be locally valid. Formally, on a $N \times N$ dense grid, the energy function is defined as

$$\begin{aligned} E(\mathbf{x}) = \frac{\Gamma\theta_{rgb}}{2\pi\sigma^2} \sum_{(i,j),(i',j')} \mathbb{1}[\mathbf{x}_{(i,j)} \neq \mathbf{x}_{(i',j')}] e^{-\frac{\|(i,j)-(i',j')\|^2}{2\sigma}} \\ + \sum_{(i,j)} U_{(i,j)} \mathbb{1}[\mathbf{x}_{(i,j)} = 0] , \end{aligned}$$

where σ controls the range of the correlations and $U_{(i,j)}$ is a unary potential.

Since that we assumed that all the variables receive the same unary U, all the variables are undiscernibles. Furthermore, the pairwise potentials are attractive, we therefore expect all the mean-field parameters $q_{i,j} = Q(\mathbf{x}_{i,j} = 0)$ to have the same value at the fixed point solution of the Mean-Field. Therefore, we designate this common parameter q^T and we can try to find analytically the Mean-Field fixed point for q^T corresponding to a temperature T.

At convergence, the parameter q^T will have to satisfy

$$\begin{aligned}\log(q^T) &= \mathbb{E}_Q(E(\mathbf{x})|x_i = 0) \\ &= -\frac{\Gamma\theta_{rgb}}{2\pi\sigma^2 T}\sum_{(i,j)\in\mathbb{Z}\times\mathbb{Z}}(1-q^T)e^{-\frac{\|(i,j)\|^2}{2\sigma}} - \frac{U}{T} \\ &= -\frac{(1-q^T)\Gamma\theta_{rgb}+U}{T}\end{aligned}$$

Hence, we obtain the fixed point equation

$$\tilde{q}^T = \frac{1}{2}\tanh\left(\frac{\tilde{q}^T\Gamma\theta_{rgb}-U}{T}\right) , \tag{2}$$

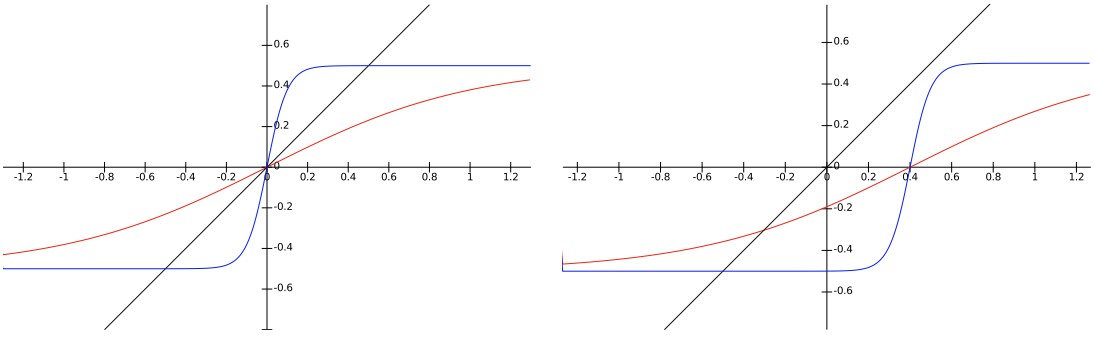
where $\tilde{q}^T = q^T - 0.5$. As depicted in Figure 1, when unaries are 0 (on the left) there are two distinct regimes for the solutions of this equation. For high T, there is only one stable solution at $\tilde{q} = 0$. For low T, there are two distinct stable solutions where $\tilde{q}$ is close to -0.5 or 0.5. The temperature threshold T_c where the transition happens, corresponds to the solution of

$$\frac{1}{2}\frac{d\tanh(\frac{\tilde{q}\Gamma\theta_{rgb}}{T})}{d\tilde{q}}|_{\tilde{q}=0} = 1 , \tag{3}$$

and hence $T_c = \frac{\Gamma\theta_{rgb}}{2}$. For real images, we have $\theta_{rgb} \leq 1$, and therefore, $T_c = \frac{\Gamma}{2}$ can be used to upper-bound the true critical temperature.

When unaries are non-zero, there is no closed form solution for T_c, however, from Equation 2, we can show that the smaller the unaries (U), the lower the critical temperature will be. This is intuitively justified in Fig. 1.

The authors of [3], use several heuristics which basically consist in looking for high correlations and low unaries directly in the potentials of the graphical model, in order to find good variables to clamp. We, instead use a criterium based on the critical temperature in order to spot these.



Without unaries With unaries

Figure 1: $\tanh(\frac{\tilde{q}\Gamma - U}{T})$ for two temperatures. Low T (blue) and High T (red).

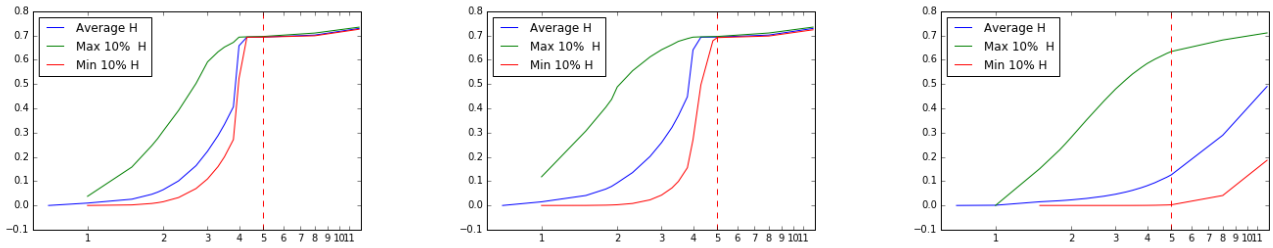
2 Experimental analysis

We use the dense CRF implementation of [2] to verify the phase transition experimentally for $\Gamma = 10$. In our experiments, we used the three following settings, which range from the stylised example used for calculation to real semantic segmentation problems:

- **Model 1:** We use a uniform rgb image $d_{rgb} = 0$. Two classes without unary potentials. This is exactly the model used for the derivations with $\theta_{rgb} = 1$ and $U = 0$.
- **Model 2:** Gaussian potentials defined over image coordinates distance + RGB distance. Two classes without unary potentials. In other words, $\theta_{rgb} \leq 1$.
- **Model 3:** Gaussian potentials defined over image coordinates distance + RGB distance. Two classes with unary potentials produced by a CNN. This is a real-life scenario.

Fig. 2 shows that, as expected, two regimes appear for Model 1, before and after $T = 5$. We see that our prediction remains completely valid for Model 2, some non-uniform regions fall under the regime $\theta_{rgb} \leq 1$ and therefore the 10 % highest entropy percentile transitions slightly earlier. For Model 3, however, we see that the minimal and average entropy remain low even for $T > 5$. This is well explained by the fact that large regions of the image receive strong unary potentials from one class or the other, and therefore fall under the case "with unaries" of Fig. 1 where the U parameter cannot be ignored. However, some uncertain regions receive unary potentials of same value for both labels, and therefore undergo a phase transition as predicted by our calculation. That is why the maximal entropy behaves similarly to Model 2. Our algorithm precisely targets these uncertain regions.

Interestingly, we see that in practice, the users of DenseCRF choose the Γ and T parameters in order to be in a Multi-Modal regime, but close to the phase transition. For instance in the public releases of [1] and [4], the Gaussian kernel is set with $T = 1$ and $\Gamma = 3$.



Model 1 Model 2 Model 3

Figure 2: Entropy as a function of temperature.

References

[1] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. Yuille. Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs. In *International Conference for Learning Representations*, 2015. 3

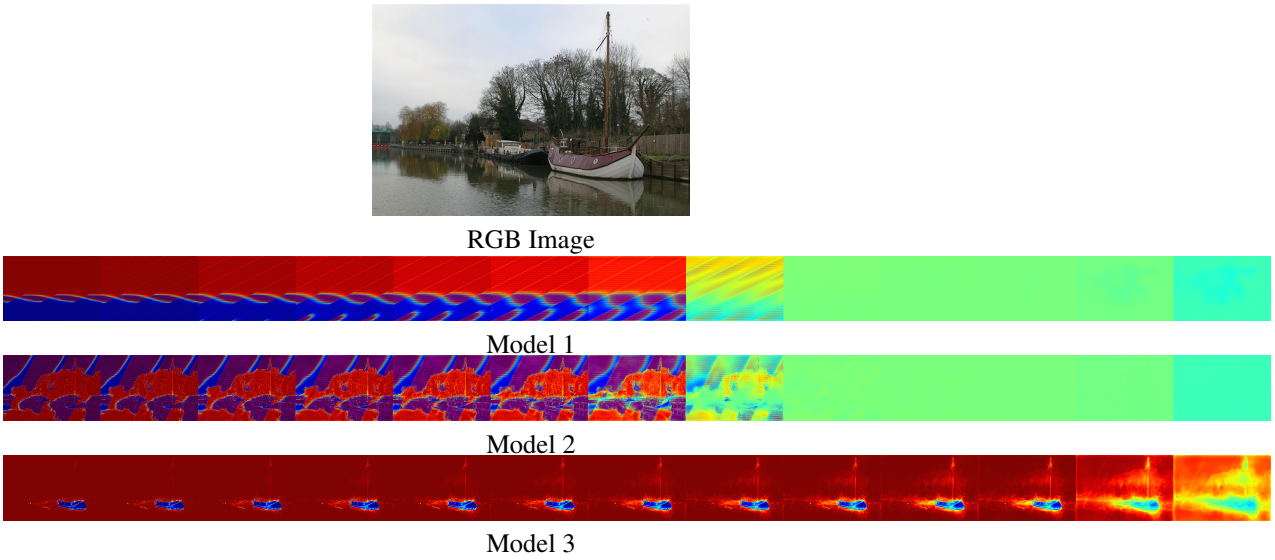


Figure 3: Evolution of MF probability for background label when temperature increases

[2] P. Krähenbühl and V. Koltun. Parameter Learning and Convergent Inference for Dense Random Fields. In *International Conference on Machine Learning*, pages 513–521, 2013. 1, 3

[3] A. Weller and J. Domke. Clamping improves trw and mean field approximations. In *Advances in Neural Information Processing Systems*, 2015. 2

[4] S. Zheng, S. Jayasumana, B. Romera-paredes, V. Vineet, Z. Su, D. Du, C. Huang, and P. Torr. Conditional Random Fields as Recurrent Neural Networks. In *International Conference on Computer Vision*, 2015. 3