noise π	max-marg.	std	mean-marginals	std	SVM-Struct	std
1%	0.4%	<0.1%	0.4%	<0.1%	0.6%	<0.1%
5%	1.1%	<0.1%	1.1%	<0.1%	1.5%	<0.1%
10%	2.1%	<0.1%	2.0%	<0.1%	2.8%	0.3%
20%	4.2%	<0.1%	4.1%	<0.1%	6.0%	0.6%

Table 1: Supervised denoising results.

	π is fixed				π is not fixed			
π	max-marg.	std	mean-marg.	std	max-marg.	std	mean-marg.	std
1%	0.5%	<0.1%	0.5%	<0.1%	1.0%	-	1.0%	-
5%	0.9%	0.1%	1.0%	0.1%	3.5%	0.9%	3.6%	0.8%
10%	1.9%	0.4%	2.1%	0.4%	6.8%	2.2%	7.0%	2.0%
20%	5.3%	2.0%	6.0%	2.0%	20.0%	-	20.0%	-

Table 2: Unsupervised denoising results.

parameter for t, one for α, both learned by cross-validation. Given our estimates, we may denoise a new image by computing the "max-marginal", e.g., the maximum a posteriori $\max_x p(x|z, \alpha, t)$ through a single graph-cut, or computing "mean-marginals" with 100 logistic samples. To calculate the error we use the normalized Hamming distance and 100 test images.

Results are presented in Table 1, where we compare the two types of decoding, as well as a structured output SVM (SVM-Struct [22]) applied to the same problem. Results are reported in proportion of correct pixels. We see that the probabilistic models here slightly outperform the max-margin formulation[1] and that using mean-marginals (which is optimal given our loss measure) lead to slightly better performance.

Unsupervised image denoising. We now only consider $N = 100$ noisy images $z_1, \dots, z_N$ to learn the model, without the original images, and we use the latent model from Section 4.4. We apply stochastic subgradient descent for the difference of the two convex functions A_{logistic} to learn the model parameters and use fixed regularization parameters equal to 10^{-2}.

We consider two situations, with a known noise-level π or with learning it together with α and t. The error was calculated using either max-marginals and mean-marginals. Note that here, structured-output SVMs cannot be used because there is no supervision. Results are reported in Table 2. One explanation for a better performance for max-marginals in this case is that the unsupervised approach tends to oversmooth the outcome and max-marginals correct this a bit.

When the noise level is known, the performance compared to supervised learning is not degraded much, showing the ability of the probabilistic models to perform parameter estimation with missing data. When the noise level is unknown and learned as well, results are worse, still better than a trivial answer for moderate levels of noise (5% and 10%) but not better than outputting the noisy image for extreme levels (1% and 20%). In challenging fully unsupervised case the standard deviation is up to 2.2% (which shows that our results are statistically significant).

6 Conclusion

In this paper, we have presented how approximate inference based on stochastic gradient and "perturb-and-MAP" ideas could be used to learn parameters of log-supermodular models, allowing to benefit from the versatility of probabilistic modelling, in particular in terms of parameter estimation with missing data. While our experiments have focused on simple binary image denoising, exploring larger-scale applications in computer vision (such as done by [24, 21]) should also show the benefits of mixing probabilistic modelling and submodular functions.

Acknowledgements. We acknowledge support the European Union's H2020 Framework Programme (H2020-MSCA-ITN-2014) under grant agreement n°642685 MacSeNet, and thank Sesh Kumar, Anastasia Podosinnikova and Anton Osokin for interesting discussions related to this work.

[1] [9] shows a stronger difference, which we believe (after consulting with authors) is due to lack of convergence for the iterative algorithm solving the max-margin formulation.