Most of Machine Learning deals with convex losses and convex regularizations, because strictly convex loss functions have a unique global minimum. Convex losses are prone to outliers and counter examples were discussed in Lecture 10 in the domain of Boosting and Regression.

We have begun the investigation of non-convex losses. The method is to design a family of non-convex losses in which the degree of non-convexity (the bending down of the wings of the convex loss) can be tuned. In a practical setting small amounts of non-convexity

- donnot seem to introduce local minima
- and yet can make the learning problem robust to outliers.

We designed three tunable loss functions:

1. Mismatched loss 1, where we replace the log_\text{t} in the matching loss for t-logistic regression with the natural logarithm ln.
2. Mismatched loss 2, where we replace ln in the logistic loss with log_\text{t}.
3. and an ”Energy based method”, where we replace the linear activations $\hat{a}_c$ in logistic regression with an S-shaped function $s(\hat{a}_c)$. This is outlined in the following paper fragment. The Energy based methods are discussed in a larger machine learning context in this tutorial.

Your job:

- Find multiclass (3 classes suffice) counter examples similar to the ones discussed in Lecture 10 that exemplify the robustness of the new methods. I suggest to start with methods 2 and then 3. Method 1 might be too hard to implement. Do that one only if you have time.
- Always compare to logistic regression.
- Here is the deep question: Explore why method 2 is better than 3.
- Either argue with graphical visualizations or with experiments on artificial data.
- Start with something simple and then build up, trusting your intuition. Clearly write up what you found.