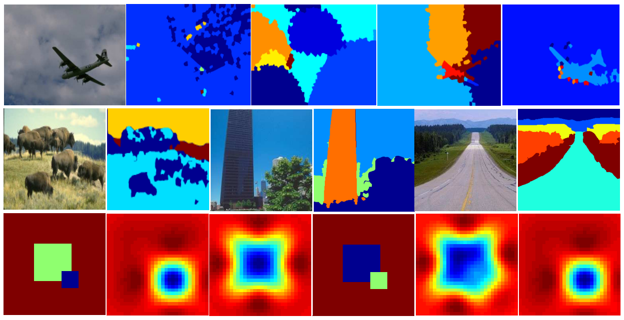

Figure 2. **Model Properties**. *TOP*- Prior samples from models employing heuristic distance+pb [25], learned distance (PY-dist) , learned distance+pb and all cues (PYall) based covariances. *CENTER*- Layered segmentations produced by our method. *BOTTOM* - Three layer synthetic partitions illustrating preferred layer orderings, Layer 1 is displayed in blue and Layer 2 in green. *Left to right:* Partition 1 (*blue = low; red = high*), the inferred Gaussian function for layers 1 and 2, partition 2 and the corresponding Gaussian functions. Under our model, partition 1 has a log probability of -77 while partition 2 has a log probability of -90.

Prior samples. Our model defines a distribution over image partitions, which can be partially assessed by visualizing partitions sampled from the prior. Figure 2 displays such samples. Note that the samples from the conditionally specified models better reflect the structure of the image.

Layers. Our model produces partitions made up of layers, not segments. These layers can have multiple connected components, due to either occlusion by a foreground layer, or a layer support function with multimodal shape. The inferred partitions illustrated in the second row of figure 2 illustrate this point. The model groups all buffaloes (in the first image), non-contiguous portions of sky, grass and trees (in the second and third images) in the same layer. Traditional segmentation algorithms, having no notion of layers, would assign each non contiguous region to a separate segment. Our layered representation provides a higher level representation of the scene than is possible with a collection of segments, which allows us to naturally deal with complex visual phenomena such as occlusion.

Implicit prior on layer order. Recall that a partition is an ordered sequence of layers, and the likelihood of a partition is governed by the likelihood of its constituent layers. Note that reordering layers can change the set of support functions which produce those layers, which in turn makes certain orderings preferable to others. In general, our GP priors prefer simple shapes over complicated ones and hence our model prefers explaining complicated shapes via an occlusion process. Figure 2 illustrates these ideas using two synthetic partitions with the order of layers 1 and 2 flipped. The model [5] prefers the partition in the first column over the one in fourth. As can be seen from the inferred layers, partition 1 is explained by the model using simpler Gaussian functions, while partition 2 has to be explained using more complicated and hence less likely Gaussian functions.

[5]Here, we have used a squared exponential covariance kernel with length scale set to half of the partition's diagonal length.

6. Experimental Results

In this section we present quantitative evaluations of various aspects of the proposed model along with qualitative results. In all experiments, our model (PYall) used a 200 dimensional low rank representation and ran 200 discrete search iterations, with three random restarts.

Experimental Setup. We benchmark the algorithm on the Berkeley Image Segmentation Dataset (BSDS300 [12]) and a subset of of Oliva and Torralba's [16] eight natural categories dataset. We sampled the first 30 images from each of the eight categories to create a 240 image dataset.

The performance of the algorithms are quantified using the probabilistic Rand Index (*PRI*) [18], and the segmentation covering (*SegCover*) metric [2]. The partitions produced by our model are made up of layers, which may not be spatially contiguous. However, the benchmarks we evaluate on, define segments to be spatially contiguous regions. To produce these we run connected components on the layers splitting them into spatially contiguous segments.

Quantifying model enhancements. This paper improves on both the model (PYheur) and the corresponding inference algorithm presented in [25] . To quantify the performance gains solely from model enhancements we devise the following test. On BSDS300 test images, we compare the log-posterior assigned to the ground truth human segmentations $p(z_{gt}|x, \eta)$ under both models. Since, we already have access to z_{gt} no inference is required and the model which assigns higher probability mass to the ground truth, models the data better. Figure 3 presents a scatter plot comparing both models. It is easy to see that PYall models human segmentations significantly better.

Evaluating inference enhancements. Next, we evaluate the performance improvements resulting from the novel inference algorithm[6]. Figure 3 displays the result of running mean field and search based inference from 10 random initializations for a given test image. The log-likelihood plots clearly demonstrate mean field being susceptible to local minima. In contrast, EP based search exhibits robustness and all chains converge to high probability partitions. The bottom row displays the best and worst partitions found by mean field and search. As one would expect, there is wide variability in the quality of mean field partitions, while the search partitions are consistently good. The rightmost top row plot displays randomly chosen partitions from the 10 EP search runs. It demonstrates a high correlation between

[6]100 search iterations takes about 30 minutes on a standard quadcore with 4GB of ram.