Put reviewer comments in italics, so it’s easy to scan for them.

REVEWER 1:

* The significance of the results is somewhat questionable, given that the improvements are by a relatively small constant factor, and the potential for possible large improvements is not discussed.

The reviewer is correct that the improvement is only to the constant factor. Further, the degree of improvement depends on the data. However, the improvement looks (empirically on standard benchmarks) very significant for image processing applications.

* There might be other plausible approaches to expedite inference.

There might be, we agree. The virtue of our approach is that its gain may vary but it never loses to the classical algorithm. Further your comment suggested other applications.

For example, the TRW-S algorithm [Kolmogorov, PAMI, 2006] is similar to belief propagation, so our approach can be applied to it as well; TRW-S is of interest because it has a high inference quality compared to other approach as shown in a survey paper [Kappes, CVPR, 2013]. We appreciate the interesting comment.

- The presentation needs a revision but is acceptable.

We appreciate your elaborative advice on increasing the readability of our paper. We have revised the paper according to the advice.

REVEWER 2:

- Giving an intuition of what the three cases in Definition (1) mean / refer to would be appreciate.

The three cases corresponds to Lemma 1, 2, and 3. We have added intuitions to explain the three cases.

- It would be interesting to know how and if the proofs would hold in graphs with more connections, or even fully connected graphs.

Although we assumed the pairwise model, we can handle fully connected graphs by slightly modifying our approach. We have added a detailed discussion.

- Reasons for why not all messages will need updating would be very much appreciated.

This is because most messages reach convergence in a first few iterations. We will add the number of converged messages and the discussion.

- The experiments are rather restricted as they only include six images.

We will show results for more datasets for which we confirmed the effectiveness of our approach. In addition, we will show actual labeling result of our and existing approaches.

REVEWER 3:

- For messages in the negative log (i.e. energy) domain (min-sum, as used in the paper), the corresponding 'normalization' is the addition of a constant to all messages.

We have revised the paper based on the comment. Note that our approach guarantees the same labeling result as the standard approach even if it employs another normalization. The reason is that our message update approach is independent from the normalization as shown in Definition 1.

- I \_really\_ dislike the name of the paper, which suggests a general algorithm for improving belief propagation, when the algorithm detailed only works for… This is quite a specific setting and does not deserve an even more general name.

We appreciate your constructive comment. We will change the name of the paper to Faster Belief Propagation for Images and Related Applications.

- The algorithm used is min-sum, not max-product as stated in the paper.

We have revised the paper based on the comment.