NIPS 2018 Reply to Reviewers

Anonymous Author(s)
Affiliation
Address
email

1 Reviewer 1

Q: What is the motivation for being Bayesian about the LISTA algorithm. How does in the uncertainty in the Bayesian LISTA procedure relate to the original linear model.

A: The LISTA algorithm introduces the weights in the model. Their point estimates are being inferred during the training phase, this leads to the uncertainty of the estimation. In our Bayesian model we infer the posterior distributions over these weights to estimate this uncertainty.

Q: how does the sparse promoting prior interact with the sparsity promoting non-linearity.

A: In the LISTA model the sequential estimates of beta are obtained at each layer using the previous estimate. At every layer the estimate becomes more sparse due to soft-thresholding non-linearity. We do not specify sparsity priors over the $beta_l$ - we show that the distribution of these estimates can be approximated with the spike-and-slab distribution, i.e. its sparsity naturally arises from the Gaussian prior on NN weights and sparsity promoting soft thresholding nonlinearity.

Q: The approximations in section 4 are part of the approximate inference

A: Yes, they are part of the inference procedure, the model remains the same as in section 3. These approximations are similar to mean-field approximations in variational inference.

Q: The assumption of mutual independence over the elements of S and $beta_{l-1}$

A: Generally speaking, correlation is partially removed due to the soft-thresholding on the previous layer. It requires further research to establish the influence of remaining correlation on the quality of approximation.

28 Reviewer 3