

---

# NIPS 2018 Reply to Reviewers

---

Anonymous Author(s)

Affiliation

Address

email

## 1 Reviewer 1

2 Q: What is the motivation for being Bayesian about the LISTA algorithm. How does in the uncer-  
3 tainty in the Bayesian LISTA procedure relate to the original linear model.

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

7  
8

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

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

15  
16

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

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

20  
21

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

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

26  
27

## 28 Reviewer 3