

Supplemental: Bayesian MRI reconstruction with structured uncertainty distributions

A Implementation details

Here we outline some implementation details for our method. Full details can be found in the code https://github.com/teojd/supn_variational_mri.

For the SUPN prior distribution we define encoder $p(z|x) = \mathcal{N}(\mu_\phi(x), \sigma_\phi^2(x)I)$ and decoder $p(x|z) = \mathcal{N}(\mu_\theta(z), \Sigma_\theta(z))$. Our parameterisation of Σ_θ is defined by a sparse representation of the Cholesky of the precision matrix as described in the original structured uncertainty prediction networks (SUPN) paper [4]. Here the original formulation has also been expanded to account for the multi channel requirement of complex MRI by adding an additional variable per pixel that captures cross-channel off-diagonal precision elements. Our architectures are composed of 5 encoding/decoding blocks, consisting of 3 residual layers similar to the original work. We also use group normalisation and SiLU activation functions. We initialise the predicted covariance to be equivalent to a scaled diagonal, $\Sigma_\theta = \frac{1}{10}I$ and then we train the prior in two stages: In stage one we fix Σ_θ and train the encoder and mean decoder for 200 epochs with a learning rate of 10^{-3} to maximise the standard VAE evidence lower bound

$$\mathcal{L}_{VAE}(\theta, \phi) := \mathbb{E}_{q_\phi(z|x)}[p(x|z)] - D_{KL}(q(\cdot|x)||p(\cdot|x)). \quad (1)$$

In stage two we reduce the learning rate to 10^{-4} , unfreeze the covariance Σ_θ and optimise the same loss, eq. (1) over all encoder and decoder parameters for a further 400 epochs.

Given this prior, posterior reconstruction is performed by maximising the evidence lower bound as described in the main paper.

B Further visualisations: True vs modelled residuals

We validate the spatial structure captured by our posterior visually by comparing samples of our posterior residual to the true residual attained by subtracting the ground truth from the posterior estimate in Fig. 1. We see for these examples that the true and posterior patterns share similar characteristics at all mask sizes. We wouldn't expect the model residual to perfectly reproduce fine detail in cases where insufficient data is provided, however we do expect the fine details of the true residuals to generally appear plausible under the posterior, which is indeed the case. Moreover, as data increases, the breakdown of the true residual structure is matched by our posterior model, demonstrating that these residuals provide visual indications of the localised spatial accuracy of a reconstruction that can not be reproduced by pixelwise or sample based uncertainty approaches.

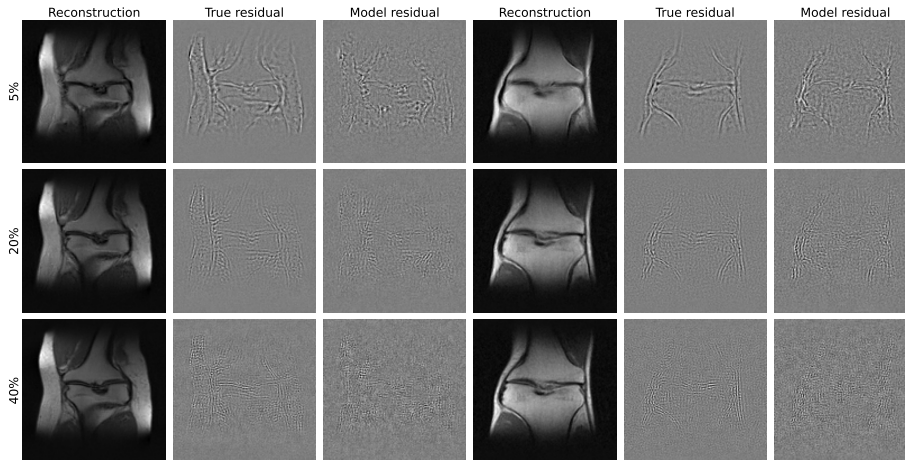


Fig. 1. Visualisation of the true and modelled residuals for two test examples at various data quantities. Our method captures the qualitative spatial features of the true residual in all cases, giving insight into the space of possible reconstruction errors without requiring ground truth.

C Variational approximation structure

We chose a separable variational approximation $q_{\lambda}(x, z) = q_{\lambda_x}(x)q_{\lambda_z}(z)$, which may appear restrictive at first since this approximate posterior does not explicitly model the correlation between x and z . This is a nuanced point since the true posterior $p(x, z|y)$ does contain correlations between x and z . However, the role of z within the framework is solely to enable flexible prior modelling, and it will ultimately be marginalised from the posterior, as we seek to estimate $p(x|y) = \int p(x, z|y)dz$. The specific values that z takes, and their correlations to x , don't provide any additional insight over the reconstruction. So our approximate posterior not explicitly modelling the correlation in the joint posterior distribution is not practically restrictive, as it still benefits from the expressive prior model over x that the introduction of z enables. The result is the best structured posterior approximation to the true posterior under this expressive prior. This choice also trivialises the marginalisation of z , and the sparse Gaussian construction ensures that we can easily evaluate posterior statistics and produce interpretable uncertainty visualisations as shown in the main paper.

Of course, constructions that do consider the conditional structure could be applied, for example by allowing our variational distribution to model images conditional on the latent by taking $q_{\lambda}(x, z) = q_{\lambda_x}(x|z)q_{\lambda_z}(z)$, with $q_{\lambda_x}(x|z)$ modelled by some decoding distribution. This approach could lead to a more flexible class of posterior, but would bring back the need for sample-based inference and ultimately harm the interpretability provided by our method.