Extended and Unscented Kitchen Sinks
Supplementary Material

Edwin V. Bonilla   E.BONILLA@UNSW.EDU.AU

The University of New South Wales

Daniel Steinberg   DANIEL.STEINBERG@NICTA.COM.AU

NICTA

Alistair Reid   ALISTAIR.REID@NICTA.COM.AU

NICTA

1. Experiments

In this section we present additional results and details of the experiments.

1.1. Synthetic inversion problem

Figure 1 shows the performance of the algorithms in terms of the SMSE on the target value ($\bar{\mathbf{g}}_*$) and the NLPD. As in the main paper, we can see that on these benchmarks the performance of the algorithms is fairly robust as a function of the number of features, with the UKS being more sensitive to a very low number of features ($D = 50$).

1.2. Seismic inversion

For the baseline MCMC, the 113 locations ($\mathbf{X}$) in the application dataset require 904 inputs ($\mathbf{f}$), which was found to be impractically large for it to converge. Therefore, for MCMC, the dimensionality of the input space was reduced by parameterizing each latent input task with 20 regularly spaced truncated cosine basis functions, and interpolating the full $\mathbf{X}$. With this set-up, approximately 10^6 samples were needed for MCMC to converge using a Gaussian proposal function. The prior over the basis weight parameters was independent Gaussian (spatial smoothness was encoded in the basis functions). The depth standard deviation was equal to 15% of the layer mean, and the velocity equal to 10% of the layer mean. The layer mean depths were 200m, 500m, 1600m and 2200m. The corresponding velocity means were 1950m/s, 2300m/s, 2750m/s and 3650m/s respectively.

The observation likelihood model used was a conditionally independent Gaussian distribution given the $\mathbf{f}$ parameters - the mean was computed using the sensor forward model, and an observation noise variance for each task (layer) was learned. The prior for this noise was a Gamma distribution with mode noise variances of 0.025, 0.05, 0.075 and 0.1 s^2 for each layer, and a mean of 1.2 times the mode.

The EKS and UKS methods are able to handle the full input dimensionality tractably, so they were applied to the full problem input space. An equivalent prior on the weight space was derived to closely match the MCMC prior (although not identical because of the higher degrees of freedom). An identical observation likelihood was used.

Additional results: Figures 2 and 3 show the predicted (mean) depths and (mean) velocities by MCMC and the EKS on the seismic experiment. We see that the EKS attains relatively accurate mean predictions for depths and velocities across layers. Furthermore, Figure 4 shows the relative errors in the predictive variances of the EKS when compared to MCMC. We see that the EKS consistently over-estimates the uncertainty but it is close to the values obtained with MCMC.
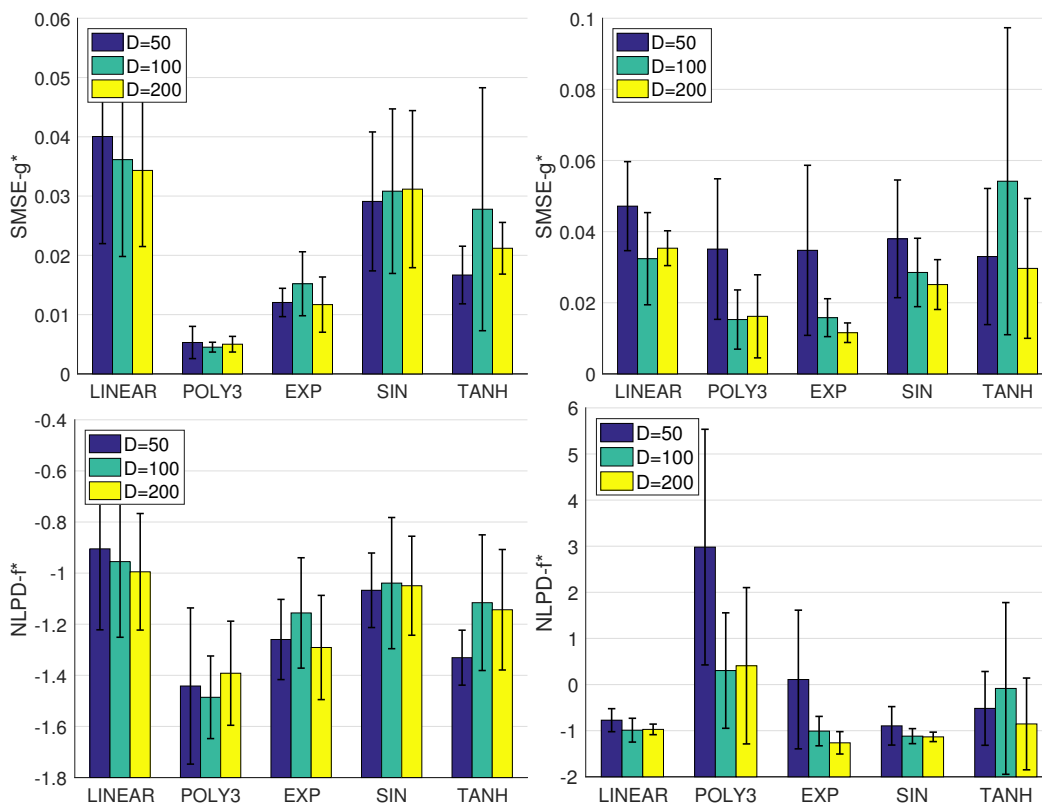
Figure 1. The performance of the EKS (left) and UKS (right) on the synthetic inversion problems as a function of the number of features.
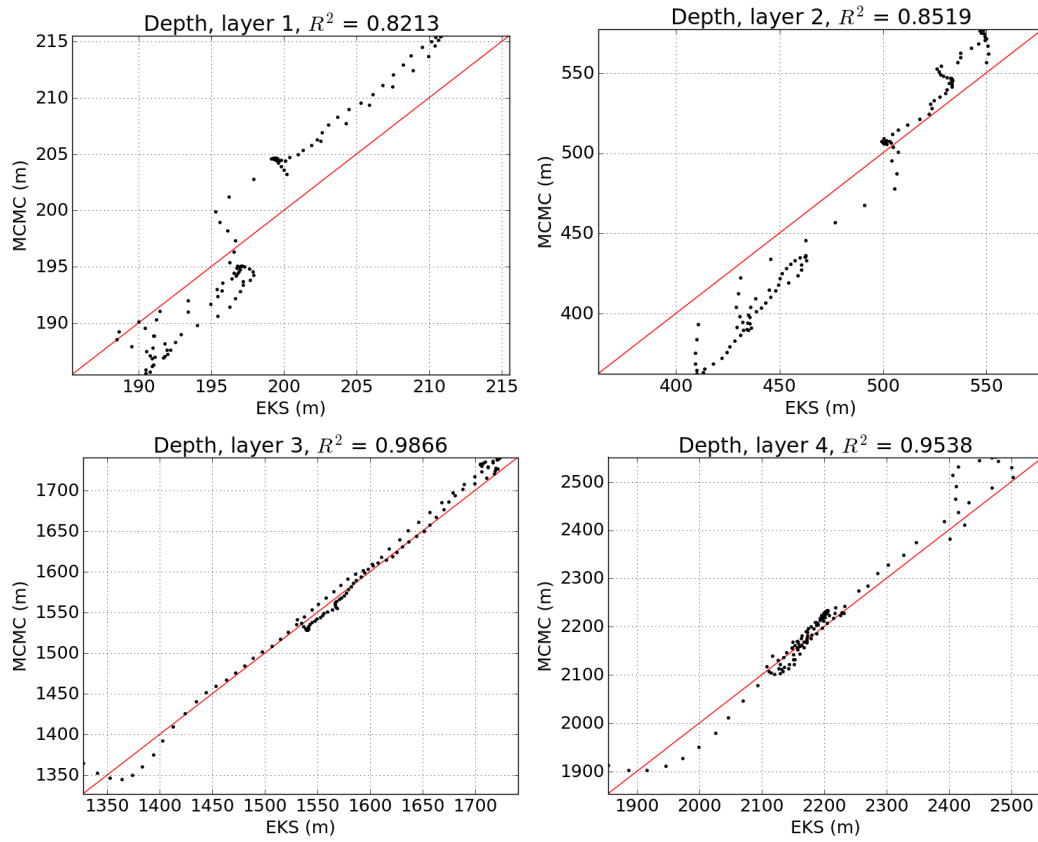
Figure 2. Scatter plots of the predicted depths of each layer in the seismic experiment by MCMC and the EKS, where R^2 denotes the coefficient of determination.
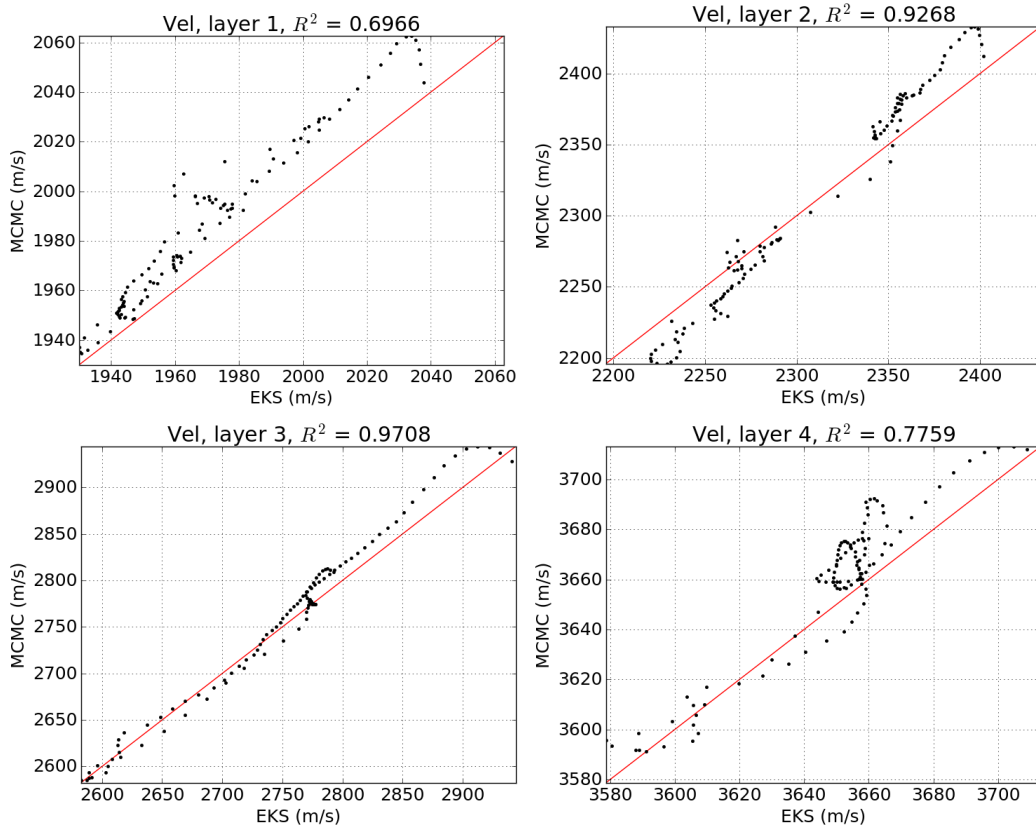
Figure 3. Scatter plots of the predicted velocities of each layer in the seismic experiment by MCMC and the EKS, where R^2 denotes the coefficient of determination.
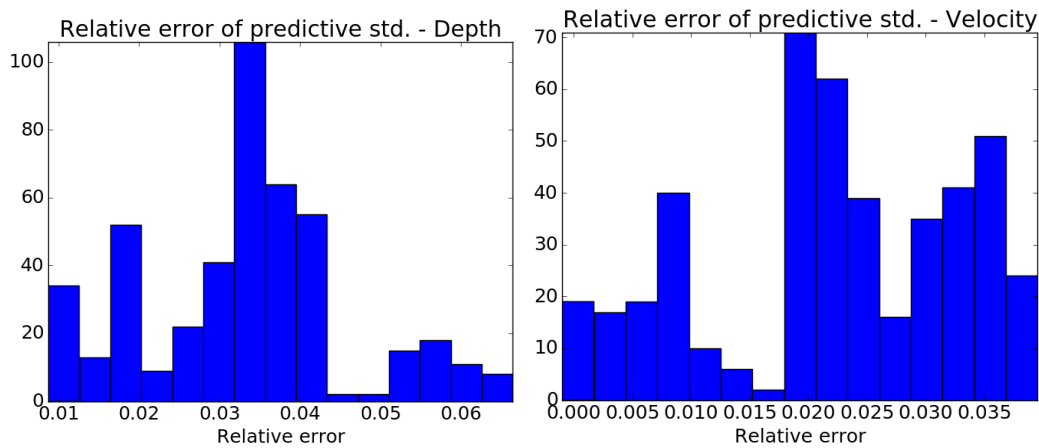

Figure 4. Histograms of the relative error in standard deviations for the EKS computed as $(\sigma_{\text{EKS}} - \sigma_{\text{MCMC}})/\mu_{\text{MCMC}}$, where μ and σ are the corresponding mean and standard deviations of the predictive distributions.