MCMC methods somewhat unattractive.

3.1 A Hierarchical Non-Gaussian Linear Model

Consider this motivating example of a linear hierarchical model discussed by Papaspiliopoulous and Roberts (2008).

$$Y = X + \epsilon_1 \tag{13}$$

$$X = \Theta + \epsilon_2 \tag{14}$$

For an observed value Y , X is the latent mean for the prior on Y, Θ is the prior mean of X, and ϵ_1 and ϵ_2 are random error terms, each with mean 0. Papaspiliopoulous and Roberts note that to improve the robustness of inference on X to outliers of Y, it is common to model ϵ_1 as having heavier tails than ϵ_2 . Let $\epsilon_1 \sim \text{Cauchy}(0, 1)$, $\epsilon_2 \sim N(0, 5)$, and $\Theta \sim N(0, 50000)$, and suppose there is only one observation available, $Y = 0$. The posterior joint distribution of X and Θ is given in Figure 1; the contours represent the logs of the computed posterior densities. Note that around the mode, X and Θ appear uncorrelated, but in the tails they are highly dependent. Papaspiliopoulous and Roberts present this example as a deceptively simple case in which Gibbs sampling performs extremely poorly. Indeed, they note that almost all diagnostic tests will erroneously conclude that the chain has converged. The reason for this failure is that the MCMC chains are attracted to, and get “stuck” in, the modal region where the variables are uncorrelated. Once the chain enters the tails, where the variables are more correlated, the chains moves slowly, or not at all.

GDS is a more effective alternative for sampling from the posterior distribution. The posterior mode and Hessian of the log posterior at the mode, are $\theta^* = (0, 0)$ and $H =$