**Horseshoe Prior**

* Based on multivariate normal scale mixtures
  + Mean parameter has prior of , where the *local* shrinkage parameters themselves have priors with a *global* shrinkage parameter that depends on the variance of the observations (uncertain if this variance is indeed determined by the data or by other methods)
  + *No user-chosen hyperparameters*
  + is the *shrinkage coefficient*
    - 🡪 1 : total shrinkage, noise
    - 🡪 0 : no shrinkage, signal
* Main advantages of horseshoe prior:
  + ***Robust to large signals***:
    - Its heavy tails mean that the horseshoe prior does not suppress large signals, does not shrink them as it would very small signals
    - 🡪 obvious signals unshrunk
  + ***Super-efficient when true answer is sparse (noise):***
    - Pole at origin (i.e. prior density is unbounded approaching 0) means that the estimator converges super-efficiently when truth is 0/near 0, compared to other priors
  + Horseshoe performs very well compared to other priors when truth is sparse, and similarly to other priors when truth is a signal
* Can show that the estimator is when variance = 1
  + plays role similar to a *shrinkage weight*
    - 🡪 0 : noise
    - 🡪 1 : signal
    - Can’t be interpreted as a *posterior probability of inclusion*, but is very similar to one and closely lines up with it
  + Possible decision rule:
    - If , signal
    - else, noise