STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **16**, October 2026, pp 3629–3636.
Published online in International Academic Press (www.IAPress.org)

An Elliptical Slice Sampler for the Beta Prime prior

Ahmed Alhamzawi, Gorgees Shaheed Mohammad *

Department of Mathematics, College of Science, University of AL-Qadisiyah, IRAQ

Abstract Global-local type priors are theoretically ideal for variable selection. However, they suffer from difficulty to implement computationally. Since these distributions usually have sharp spikes and heavy tails, then standard sampling methods often struggle from their complex geometry. The beta prime slice sampler is introduced to solve some of these problems by solving for the lower bound to demonstrate how the prior is highly concentrated, with elliptical arcs restricted to a small region near the origin. An efficient algorithm is introduced by combining the Elliptical Slice Sampling. Simulations tests are to compare the new sampler with other established methods. It is shown that the new sampler reduces the Mean Squared Error by half compared to the Lasso. Furthermore, it offers greater stability than the Elastic Net and standard Horseshoe implementations. These results prove that the proposed method is a practical and robust solution for high-dimensional data.

Keywords Beta Prime, Slice Sampler, Metropolis-Hastings, Horseshoe, Sparsity.

AMS 2010 subject classifications 60Exx, 60E05

DOI: 10.19139/soic-2310-5070-3496

1. Introduction

The beta prime prior, also known as the Inverted Beta prior, is a powerful sampling tool that has many attractive properties. It is often used to solve a major problem in modern analysis of variable selection in high-dimensional data. In fields like genomics and finance, researchers often face a difficult challenge. The number of possible predictors (p) is frequently much larger than the number of observations (n). However, in these cases it is assumed that sparsity is present such that in situations of many variables, only a few are truly relevant. The beta prime prior can handle this issue by allowing the model to automatically select true signals from irrelevant noise [3, 4].

The standard Gaussian linear regression model is:

$$\mathbf{Y} = \mathbf{X}\beta + \epsilon, \tag{1}$$

where $\mathbf{y} \in \mathbb{R}^n$ is the vector of response variables, $\mathbf{X} \in \mathbb{R}^{n\times p}$ is the design matrix of predictors, $\beta \in \mathbb{R}^p$ is the vector of regression coefficients and $\epsilon \in \mathbb{R}^n$ is the vector of error terms, with $\epsilon \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$.

In Bayesian Linear Regression, variable selection depends heavily on which prior distribution is chosen for the coefficients (β) to represent initial assumptions about the size of these coefficients before data is observed. A simple, but limited prior is the Ridge prior [5] which works by shrinking all coefficients toward zero. However, this method applies pressure uniformly. Its main flaw is that it cannot separate true signals from random noise. It tends to shrink important coefficients too much. Meanwhile, it usually fails to force noise coefficients all the way to zero which makes the model difficult to interpret [6]. To deal with sparse situations we need a better solution. Some authors have led to the development of global-local shrinkage priors like the beta prime prior [7]. These methods

*Correspondence to: Gorgees Shaheed Mohammad (Email: gorgees.alsalamy@qu.edu.iq).Department of Mathematics, College of Science, University of AL-Qadisiyah, IRAQ.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

use a global parameter to manage the overall level of sparsity and a local parameter to allow specific coefficients to control the distance from the mean. If the data shows a strong signal, then the local parameter increases. Thus, this prevents the coefficient from shrinking too much. Alternatively, the local parameter stays small for noise variables, allowing the global parameter to aggressively push them towards zero [8].

We can see that this hierarchy presents a marginal prior characterized by a sharp peak at zero with heavy, Cauchy-like tails [8, 9, 10]. The shape of this prior is important because it solves the bias-variance trade-off problem in traditional penalization methods such as the Lasso since it often overshrinks true coefficients in order to suppress the noise [11, 12]. By possessing heavy tails, global-local priors ensure that large signals remain essentially unbounded and once the local parameter inflates in response to strong data evidence, it effectively neutralizes the global shrinkage pressure [13]. This property allows the model to recover sparse signals with near-unbiased estimates. Hence, it serves as a computationally efficient and continuous approximation to the ideal discrete "spike-and-slab" variable selection methods [14, 15].

Unfortunately, the beta prime prior has a computational cost since the posterior distribution cannot be solved directly. To use this model, we rely on Markov Chain Monte Carlo (MCMC) methods. Specifically, we employ a Gibbs Sampler [16]. This approach breaks the complex problem down into simpler updates. However, standard methods often struggle with shrinkage priors because they fail to handle the sharp spikes and heavy tails effectively. To fix this, we aim to insert the slice sampler within the Gibbs steps [17]. By combining Gibbs and Slice sampling, we can ensure the algorithm works efficiently. It navigates the difficult areas where other methods can get stuck, making variable selection feasible [18]. The integration of Slice Sampling into the Gibbs framework is an attractive proposition. It unifies Bayesian computation significantly. This approach inverts the standard logic of Elliptical Slice Sampling. Consequently, the algorithm becomes independent of the prior's specific shape. It also simplifies high-dimensional inference and the benefits of adaptive sparsity can be applied across many scenarios [18]. In this paper we will introduce the beta prime slice sampler to address some of these problems. This efficient algorithm will combine the Elliptical Slice Sampling benefits with the attractive properties of the beta density. We will use simulation tests to compare the new sampler with other established methods to show how the new sampler reduces the mean square error while offering greater stability than other methods.

2. The Beta Prime prior

Our aim is to build an algorithm that provides an efficient Markov Chain Monte Carlo (MCMC) method to sample the posterior distribution of the Gaussian linear regression model with a beta prime prior. Variable selection in high-dimensional linear regression requires a specific type of prior. It must handle many small noise coefficients and a few large signal coefficients simultaneously. The beta prime prior solves this problem by using a global-local shrinkage framework [7, 19]. In this model, each coefficient follows a normal distribution: $\beta_j \sim \mathcal{N}(0, \sigma^2\tau^2\lambda_j^2)$, where the local variance λ_j^2 follows a beta prime distribution, $\lambda_j^2 \sim \mathrm{BP}(a, b)$. Its density is defined as

$$f(x) \propto x^{a-1}(1+x)^{-(a+b)}. \tag{2}$$

The two hyperparameters, a and b, govern this prior. The parameter a controls behavior near the origin such that as $a \to 0$, the prior places infinite mass near zero variance. Thus, this property leads to aggressive shrinkage of the noise variables [3]. On the other hand, the parameter b manages the shape of the tails by controlling the rate of decay $\propto (\lambda^2)^{-(b+1)}$. Hence, small values of b give heavy tails and prevent the overshrinkage of large signals while avoiding the bias commonly seen in the exponential tails of the Bayesian Lasso [8, 20]. The attractive properties of this model come with several computational problems. This is because the posterior complexity increases as the shape turns to an unbounded spike at zero and heavy tails. Some sampling algorithms such as Gibbs and Metropolis-Hastings can fail to sample efficiently in such scenarios [17]. To solve this problem, we propose to use a hybrid MCMC strategy by combining an elliptical slice sampler (ESS) for the coefficients with a standard Gibbs method for updating the scale parameters [18]. The method flips standard ESS logic by treating the Gaussian likelihood as the base measure and the beta prime prior as the likelihood target.

The beta prime prior on the regression coefficients $\beta = (\beta_1, \ldots, \beta_p)^T$ is defined hierarchically using a Gamma-Inverse-Gamma mixture. For hyperparameters $a > 0$ and $b > 0$:

$$\begin{aligned} \beta_j \mid \lambda_j^2, \sigma^2 &\sim \mathcal{N}(0, \lambda_j^2 \sigma^2) \\ \lambda_j^2 \mid \gamma_j &\sim \text{Inv} - \text{Gamma}(a, \gamma_j) \\ \gamma_j &\sim \text{Gamma}(b, 1) \\ \sigma^2 &\sim \text{Inv} - \text{Gamma}(\nu_0/2, \nu_0/2) \quad (\text{e.g., } \nu_0 = 1 \text{ or } 0) \end{aligned}$$

Note: The parameterization for the Inverse-Gamma density used here is $p(x) \propto x^{-(\alpha+1)} e^{-\beta/x}$, and for Gamma it is $p(x) \propto x^{\alpha-1} e^{-\beta x}$.

3. MCMC Algorithm

The sampler iterates between two main blocks: the multivariate update of β using the elliptical slice sampler (ESS) (Block 1), and the component-wise updates of the scale parameters using standard Gibbs sampling (Block 2). The elliptical slice sampler is a Markov Chain Monte Carlo method designed to sample from a target distribution defined by [21, 18]

$$p(\Delta) \propto \mathcal{N}(\Delta; 0, \Sigma) L(\Delta) \tag{3}$$

where Δ represents the vector of parameters of interest, $\mathcal{N}(\Delta; 0, \Sigma)$ denotes a multivariate Gaussian prior distribution with covariance matrix Σ and $L(\Delta)$ represents an arbitrary likelihood function. The algorithm updates the current value of Δ and a variable $\mathbf{v} \sim \mathcal{N}(0, \Sigma)$. The new value Δ^* is given by

$$\Delta^* = \Delta \cos\theta + \mathbf{v} \sin\theta \tag{4}$$

where $\theta \in [0, 2\pi]$ is selected using a slice sampling procedure [17, 22]. Then, another threshold variable u is drawn uniformly from the interval $[0, L(\Delta)]$. If a proposed value of θ' makes the likelihood satisfy $L(\Delta^*) > u$, then it is accepted. Otherwise, the bracket is shrunk around θ' until a valid value is found.

3.1. Initialization

Initialize $\beta^{(0)}$, $\lambda^{(0)}$, $\gamma^{(0)}$, and $\sigma^{(0)}$.

1. Update $\beta^{(t)}$: Run the elliptical slice sampler (ESS) conditional on $\sigma^{2(t-1)}$ and prior parameters.
2. Update $\sigma^{2(t)}$: Sample the error variance.
3. Update Scale Parameters: Sample $\lambda_j^{2(t)}$ and $\gamma_j^{(t)}$ for $j = 1, \ldots, p$.

3.2. Block 1: Elliptical slice sampler for β

The ESS samples the coefficient vector β by sampling its offset $\Delta = \beta - \hat{\beta}$ from the multivariate Gaussian distribution associated with the data likelihood (acting as the base measure), with the beta prime prior acting as the likelihood function in the ESS framework.

Defining the parameters (at iteration t):

- OLS Estimate (Mean Vector): $\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$
- Covariance Matrix: $\mathbf{V} = \sigma^{2(t-1)}(\mathbf{X}^T\mathbf{X})^{-1}$
- Prior Log-Density $L(\cdot)$:

$$L(\beta) = \sum_{j=1}^{p} \log\left(\text{BetaPrime}(\beta_j \mid a, b)\right)$$

Alternatively, conditional on latent scales (often more stable):

$$L(\beta) = \sum_{j=1}^{p} \log\left(\mathcal{N}(\beta_j \mid 0, \lambda_j^{2(t-1)}\sigma^{2(t-1)})\right)$$

ESS Steps (Given $\beta_{curr} = \beta^{(t-1)}$):

1. Generate Proposal Direction:
$$\mathbf{v} \sim \mathcal{N}(\mathbf{0}, \mathbf{V})$$
Set the current offset $\Delta_{curr} = \beta_{curr} - \hat{\beta}$.
2. Determine Slice:
$$z = L(\beta_{curr}) + \log U, \quad \text{where } U \sim \text{Unif}(0,1)$$
3. Propose Initial Angle and Interval:
$$\theta \sim \text{Unif}(0, 2\pi)$$
Initialize angle bounds: $[\theta_{\min}, \theta_{\max}] = [\theta - 2\pi, \theta]$.
4. Iterative Slice-Shrinking: While $L(\beta_{prop}) < z$:
   - Calculate the proposed offset and vector:
$$\Delta_{prop} = \Delta_{curr}\cos(\theta) + \mathbf{v}\sin(\theta)$$
$$\beta_{prop} = \hat{\beta} + \Delta_{prop}$$
   - Shrink Interval: If $L(\beta_{prop}) < z$:
     - If $\theta < 0$, set $\theta_{\min} = \theta$.
     - Otherwise, set $\theta_{\max} = \theta$.
   - Draw a new angle $\theta \sim \text{Unif}(\theta_{\min}, \theta_{\max})$.
5. Accept: Set $\beta^{(t)} = \beta_{prop}$.

3.3. Block 2: Gibbs Updates for Scale Parameters

These updates use the conjugate relationships for the beta prime hierarchy.

1. **Update Error Variance** (σ^2)
$$\sigma^{2(t)} \mid \cdot \sim \text{Inv} - \text{Gamma}\left(\frac{n+p+\nu_0}{2}, \frac{\|\mathbf{Y} - \mathbf{X}\beta^{(t)}\|^2 + \sum_{j=1}^{p} \frac{\beta_j^{(t)2}}{\lambda_j^{2(t-1)}} + \nu_0}{2}\right) \tag{5}$$
2. **Update Local Shrinkage Parameters** (λ_j^2, γ_j) For each $j = 1, \ldots, p$:
   - Step 2a: Update Mixing Parameter γ_j
$$\gamma_j^{(t)} \mid \cdot \sim \text{Gamma}\left(a+b, 1 + \frac{1}{\lambda_j^{2(t-1)}}\right) \tag{6}$$
   - Step 2b: Update Variance Scale λ_j^2
$$\lambda_j^{2(t)} \mid \cdot \sim \text{Inv} - \text{Gamma}\left(a + \frac{1}{2}, \gamma_j^{(t)} + \frac{\beta_j^{(t)2}}{2\sigma^{2(t)}}\right) \tag{7}$$

3.4. Beta Prime prior lower bound

The beta prime distribution arises when we consider the distribution of the squared local scale parameter, $Y = \lambda^2$. The beta prime distribution is also known as the Inverted Beta distribution.

Consider the probability density function $f(\theta)$ defined for any $a, b > 0$:

$$f(\theta) = \frac{\Gamma(b+1/2)}{\sqrt{2\pi}B(a,b)} U\left(b+\frac{1}{2}, \frac{3}{2}-a, \frac{\theta^2}{2}\right) \tag{8}$$

Following [23], we represent the Tricomi U function as:

$$U(A,C,z) = \frac{1}{\Gamma(A)} \int_0^\infty e^{-zt} t^{A-1} (1+t)^{C-A-1} dt \tag{9}$$

By substituting $A = b + 1/2$ and $C = 3/2 - a$, we obtain a specific case of the generalized hypergeometric integral. As shown in [12], for any $p > 0$ and $x > 0$, the following holds:

$$\int_0^\infty \frac{e^{-xt} t^{b-1/2}}{(1+t)^{a+b}} dt \geq \frac{\Gamma(b+1/2)}{\Gamma(a+b)} \int_0^\infty \frac{e^{-xt}}{(1+t)^{a+b}} dt \tag{10}$$

This step simplifies the kernel while preserving the singularity behavior at $x = 0$. We now apply the analytic inequality for integrals of the form $\int_0^\infty e^{-xt}(1+t)^{-p} dt$ established by [24], setting $p = a + b$ and $x = \theta^2/2$, the integral is bounded as:

$$\int_0^\infty \frac{e^{-\frac{\theta^2}{2}t}}{(1+t)^{a+b}} dt \geq \frac{1}{2(a+b)} \ln\left(1 + \frac{4(a+b)}{\theta^2}\right) \tag{11}$$

Combining the constants $C_{a,b}$ with the results, we arrive at:

$$f(\theta) \geq \frac{\Gamma(b+1/2)}{2\sqrt{2\pi}B(a,b)\Gamma(a+b+1)} \ln\left(1 + \frac{4(a+b)}{\theta^2}\right) \tag{12}$$

We will use the lower bound as an approximation for the prior in the algorithm because of the fundamental reliance on the prior in the mechanism of the algorithm. Unlike standard MCMC methods, the ESS generates a search path by forming an ellipse between the current state and a point sampled, thus if the prior is highly concentrated or has tight lower bound on the variance, then the resulting elliptical arcs are restricted to a small region near the origin.

3.5. Interpretation

Our algorithm works by iterating between firstly updating the coefficient vector β and then secondly updating the variance parameters one by one. We start by setting initial values for $\beta^{(0)}$, $\lambda^{2(0)}$, $\gamma^{(0)}$, and $\sigma^{2(0)}$. At each iteration t, we let the first block update the regression coefficients β using the ESS and define the current mean vector $\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$ and the covariance $\mathbf{V} = \sigma^{2(t-1)}(\mathbf{X}^T\mathbf{X})^{-1}$ using the data. The ESS will sample from $\mathbf{v} \sim \mathcal{N}(\mathbf{0}, \mathbf{V})$. Next, the algorithm identifies a "slice" of the beta prime prior density, where this density is calculated as $L(\beta) = \sum \log(\text{BetaPrime}(\beta_j \mid a, b))$. The slice threshold is defined as $z = L(\beta^{(t-1)}) + \log U$, where U is a random variable from 0 to 1. Using the lower bound approximation, if the prior density at this new point is higher than z, we accept the proposal. Otherwise, we shrink the search interval and pick a new point.

The second block handles the variance and scale parameters using standard Gibbs updates. This approach takes advantage of the conjugate hierarchy found in the beta prime distribution. Specifically, we update the error variance σ^2 drawn directly from an Inverse-Gamma distribution. Next, we update the local shrinkage parameters using latent variables. We first sample the variables γ_j from $\text{Gamma}(a+b, 1+1/\lambda_j^{2(t-1)})$, then update the local variances λ_j^2 using their conditional Inverse-Gamma posteriors: $\lambda_j^{2(t)} \sim \text{Inv}-\text{Gamma}(a+1/2, \gamma_j^{(t)} + \beta_j^{(t)2}/(2\sigma^{2(t)}))$. This two-block strategy is very powerful since it mixes the global moves of the ESS with efficient Gibbs sampling, thus ensuring robust convergence. This method remains effective even in high-dimensional sparse scenarios.

4. Performance Results

We aim to validate the performance of our new algorithm empirically. To achieve this, we ran tests across three different settings with increasing sample sizes and sparsity. These include the first experiment with $n = 100$ observations and $p = 8$ predictors, the second with $n = 200$ and $p = 25$, and the third representing a larger test with $n = 500$ and $p = 100$. We ran all simulations with 15,000 iterations with 3000 burn-in and 100 independent simulation runs. The comparison is between the new algorithm (BP Slice), the standard Horseshoe, the Lasso and Elastic Net. Our chosen evaluation metric is the Mean Squared Error (MSE) whereby the compared estimates to the true signal β_{true} and the standard deviation (sd) is used to assess stability.

The simulation results are summarized in Tables 1, 2, and 3. In general, they show a clear advantage for the new algorithm over other methods. In the first test $(n = 100, p = 8)$, the BP Slice achieved the lowest Mean Squared Error with a big gap in some cases. This gap confirms that the beta prime prior preserves true signals better than the Lasso, thus reducing the bias.

Interestingly, the third case $(n = 500, p = 100)$ offers the most important findings. In this large test, it proves that heavy-tailed priors are theoretically superior with robust computation. While the new method kept a tiny lead, it is clear that integrating the slice sampling gives a noticeable advantage. This shows that whether using beta prime or Horseshoe, this method beats optimization tools like Lasso in sparse regression. Furthermore, this method also shows better stability with a standard deviation of approximately 0.020, indicating reliable convergence to the true signal.

Table 1. Comparative performance (MSE and Standard Deviation) for Experiment 1 $(n = 100, p = 8)$

Method	MSE	sd
BP Slice	0.0229	0.0212
Lasso	0.0453	0.0402
Elastic Net	0.1667	0.0878
Horseshoe	0.0361	0.0325

Table 2. Comparative performance (MSE and Standard Deviation) for Experiment 2 $(n = 200, p = 25)$

Method	MSE	sd
BP Slice	0.0236	0.0192
Lasso	0.0426	0.0344
Elastic Net	0.1596	0.0897
Horseshoe	0.0378	0.0271

Table 3. Comparative performance for Experiment 3 $(n = 500, p = 100)$

Method	MSE	sd
BP Slice	0.0246	0.0199
Lasso	0.0465	0.0437
Elastic Net	0.1449	0.0908
Horseshoe	0.0274	0.0227

5. Conclusion

Variable selection in high-dimensional areas remains a central problem in statistics [1]. Our paper shows that while the prior choice is critical, the computational method used is just as important for efficiency. The beta prime prior

offers a better approach due to its ability to treat sparse and dense data. It uses a sharp spike to ignore the noise and heavy tails to protect signals from bias [3, 7]. However, the relatively complex shape of this prior often makes standard sampling tools struggle to handle the data and become inefficient [17].

In this paper, we have proposed the beta prime slice sampler that solves this computational difficulty. It achieves this by embedding the elliptical slice sampler inside a Gibbs hierarchical representation. We treated the Gaussian likelihood as the base measure and the beta prime prior as the target. This method bypasses the problem of tuning Metropolis-Hastings proposals in high dimensions [18]. The simulations in this paper provide strong empirical proof whereby the estimator cuts the MSE in half compared to the Lasso. Furthermore, it demonstrated better stability and convergence than the Elastic Net and standard Horseshoe methods. These findings show that the bias built into methods like the Lasso [6] is not a necessary price to pay for speed. Hence, this suggests that prior combined with an efficient slice sampling offers a better alternative to standard shrinkage estimators. It is highly effective for sparse regression problems in diverse fields ranging from genomics to econometrics. Future research can extend this work to non-Gaussian data, including logistic or survival models. In these areas, the beta prime prior properties could simultaneously improve prediction and variable selection.

REFERENCES

1. T. Hastie, R. Tibshirani, and M. Wainwright, *Statistical learning with sparsity: the lasso and generalizations*, CRC press, 2015.
2. G. Senn, N. Glatt-Holtz, G. Carigi, A. Holbrook, and H. Tjelmeland, *Multiproposal elliptical slice sampling*, arXiv preprint arXiv:2602.22358, 2026.
3. R. Bai and M. Ghosh, *High-dimensional multivariate posterior consistency under global-local shrinkage priors*, Journal of Multivariate Analysis, vol. 167, pp. 157–170, 2018.
4. R. Filippi, P.R. Hahn, J. He, and H.F. Lopes, *Financial literacy and perceived economic outcomes*, Journal of Applied Statistics, vol. 49, no. 10, pp. 2501–2515, 2022.
5. A.E. Hoerl and R.W. Kennard, *Ridge regression: Biased estimation for nonorthogonal problems*, Technometrics, vol. 12, no. 1, pp. 55–67, 1970.
6. R. Tibshirani, *Regression shrinkage and selection via the lasso*, Journal of the Royal Statistical Society: Series B (Methodological), vol. 58, no. 1, pp. 267–288, 1996.
7. N.G. Polson and J.G. Scott, *Shrink globally, act locally: Sparse bayesian regularization and prediction*, In Bayesian Statistics 9, pp. 501–538. Oxford University Press, 2010.
8. C.M. Carvalho, N.G. Polson, and J.G. Scott, *The horseshoe estimator for sparse signals*, Biometrika, vol. 97, no. 2, pp. 465–480, 2010.
9. N.G. Polson and J.G. Scott, *On the half-cauchy prior for a global scale parameter*, Bayesian Analysis, vol. 7, no. 4, pp. 887–902, 2012.
10. S. Takeno, Y. Inatsu, I. Karasuyama, and I. Takeuchi, *Preferential bayesian optimization with skew gaussian processes*, arXiv preprint arXiv:2305.18820, 2023.
11. S.L. van der Pas, B.J.K. Kleijn, and A.W. van der Vaart, *The horseshoe estimator: Posterior concentration around nearly black vectors*, Electronic Journal of Statistics, vol. 8, no. 2, pp. 2585–2618, 2014.
12. A. Bhadra, J. Datta, N.G. Polson, and B. Willard, *Lasso meets horseshoe: A unified approach to continuous sparse selections*, Journal of the Royal Statistical Society Series B: Statistical Methodology, vol. 81, no. 5, pp. 921–946, 2019.
13. J. Datta and J.K. Ghosh, *Asymptotic properties of bayes risk for the horseshoe prior*, Bayesian Analysis, vol. 8, no. 1, pp. 111–132, 2013.
14. H. Ishwaran and J.S. Rao, *Spike and slab variable selection: frequentist and bayesian strategies*, The Annals of Statistics, vol. 33, no. 2, pp. 730–773, 2005.
15. V. Ročková and E.I. George, *Emvs: The em approach to bayesian variable selection*, Journal of the American Statistical Association, vol. 109, no. 506, pp. 828–846, 2014.
16. S. Geman and D. Geman, *Stochastic relaxation, gibbs distributions, and the bayesian restoration of images*, IEEE Transactions on Pattern Analysis and Machine Intelligence, no. 6, pp. 721–741, 1984.
17. R.M. Neal, *Slice sampling*, The Annals of Statistics, vol. 31, no. 3, pp. 705–741, 2003.
18. P.R. Hahn, J. He, and H.F. Lopes, *Efficient sampling for gaussian linear regression with arbitrary priors*, Journal of Computational and Graphical Statistics, vol. 28, no. 1, pp. 142–154, 2018.
19. A. Alhamzawi and G.S. Mohammad, *Bayesian analysis of left censored regression with normal-compound gamma priors*, Bayesian Analysis, vol. 12, pp. 20–2024, 2024.
20. J. Johndrow, P. Orenstein, and A. Bhattacharya, *Scalable approximate mcmc algorithms for the horseshoe prior*, Journal of Machine Learning Research, vol. 21, no. 73, pp. 1–61, 2020.
21. I. Murray, R. Adams, and D. MacKay, *Elliptical slice sampling*, In Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, pp. 541–548. JMLR Workshop and Conference Proceedings, 2010.
22. Anonymous Authors, *A fast, robust elliptical slice sampling implementation for linearly truncated multivariate normal distributions*, arXiv preprint arXiv:2407.10449, 2024.
23. M. Abramowitz and I.A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55. US Government printing office, 1964.

24. H. Alzer, *On some inequalities for the incomplete gamma function*, Mathematics of Computation of the American Mathematical Society, vol. 66, no. 218, pp. 771–778, 1997.