

Partial Order Dimension Bayesian Inference with RJMCMC

1 Motivation

In many applications involving **partially ordered sets**, one often wishes to represent or learn the poset structure via multiple linear extensions (i.e., total orders) that intersect to form the target partial order. This perspective on posets is directly connected to the classical notion of **dimension**: a partial order h can be seen as the intersection of exactly K total orders if and only if $\dim(h) = K$. This dimension is a central concept in poset theory—dating back to the 1941 paper of Dushnik and Miller—and it can be surprisingly large (up to $n/2$ items).

From a partial order inference point, if we believe the observed partial order is “complex,” we may allow a higher-dimensional representation. Conversely, if we expect a “simple” poset, a smaller K may suffice. In a latent-variable or Bayesian viewpoint, one often encodes this partial order relationship via a latent matrix U , where $U \in \mathbb{R}^{n \times K}$ has K columns. After a comprehensive discussion about the dimension k , **our goal** is to employ a Bayesian framework—potentially with **Reversible Jump Markov Chain Monte Carlo (RJMCMC)**—that can discover or adaptively infer this dimension K from data (or, in the absence of data, from a prior distribution).

In what follows, we discuss:

1. **Partial order** and its dimension,
2. How **Reversible Jump MCMC** (RJ-MCMC) accommodates changes in dimension,
3. **Examples** using the Crown poset and MCMC (both fixed- K and dimension-jumping),
4. **Further implications** of dimension inference in Partial Order

2 Partial Orders and Their Dimension

2.1 The Dimension of a Partial Order

Definition (Dimension). Let $h = (X, \prec_h)$ be a partial order on a finite set X . The *dimension* of h , denoted $\dim(h)$, is the smallest integer K for which there exist K linear extensions $\ell_1, \ell_2, \dots, \ell_K$ such that

$$x \prec_h y \iff (x <_{\ell_k} y \text{ for every } k = 1, \dots, K).$$

Equivalently,

$$h = \bigcap_{k=1}^K \ell_k,$$

meaning each pair (x, y) is comparable in h if and only if it is comparable in *all* of those linear orders.

Interpretation. Geometrically, one can embed each element $x \in X$ as a point in \mathbb{R}^K , where coordinate ordering recovers the partial order. Dushnik and Miller (1941) first formalized this. Yannakakis (1982) highlighted the **computational complexity** of finding or approximating $\dim(h)$.

2.2 Critical Pairs and Realizers

Critical Pair. A *critical pair* (x, y) in a poset P is an incomparable pair that, once placed into one or another order, triggers certain transitive constraints. Intuitively, to represent every incomparable pair (x, y) correctly, one needs at least one linear extension that “flips” it (i.e., puts $y \prec x$) and another extension that puts $x \prec y$.

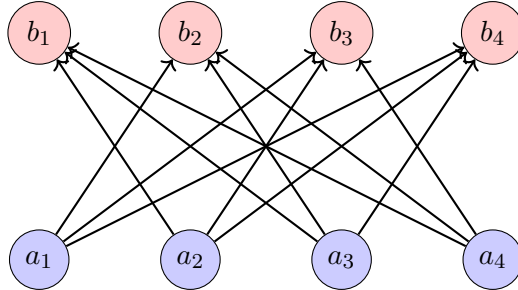
Realizer. A *realizer* of size K for $h = (X, \prec_h)$ is a family of K linear extensions whose intersection equals h . Thus, $\dim(h)$ is the size of the smallest realizer. This is directly related to how many incomparabilities (and hence critical pairs) must be “covered” by different total orders.

2.3 Crown Poset (A Standard Example)

A classic example that *maximizes* dimension for $n = 2k$ is the **crown poset** S_{2k} . It partitions $2k$ elements into two sets $A = \{a_1, \dots, a_k\}$ and $B = \{b_1, \dots, b_k\}$, with $a_i \prec b_j$ whenever $i \neq j$. The resulting partial order has:

- **Height = 2** (all edges go from A to B),
- **Dimension = k** (it attains the Hiraguchi bound $\lfloor n/2 \rfloor$).

It can be shown that $\dim(S_{2k}) = k$, and in particular for even $n = 2k$, the worst-case dimension is $\lfloor n/2 \rfloor$.



2.4 Complexity Aspects

Determining $\dim(h)$ is generally NP-hard for arbitrary posets. Yannakakis showed that even height-1 posets can have dimension inference that is NP-complete. So, from a computational perspective, **exact** dimension finding is difficult, particularly for large n . This challenge motivates approximate methods or **Bayesian** approaches that let the posterior distribution over K to infer the plausible range of dimensions.

3 Reversible Jump MCMC for Dimension K

3.1 Bayesian Model Setting

We can treat dimension K as a random variable—e.g., imposing a truncated Poisson prior $\pi(K) = \text{Pois}(\lambda)$ restricted to $K \geq 1$. For each K , the partial order is represented by an $n \times K$ latent matrix U , so that the row coordinates of each item define one or more total orders whose intersection yields h .

Latent Variables: U of Dimension K

Suppose we have a set of items

$$M = \{1, 2, \dots, n\}.$$

Each item $j \in M$ is associated with a latent vector $U_{j,:} \in \mathbb{R}^K$. Thus, the full latent matrix $U \in \mathbb{R}^{n \times K}$ has one row per item. Each row $U_{j,:}$ is an independent draw from a multivariate normal distribution:

$$U_{j,:} \sim \mathcal{N}(0, \Sigma_\rho), \quad j = 1, \dots, n,$$

where $\Sigma_\rho \in \mathbb{R}^{K \times K}$ is the covariance matrix governing the correlations between the K latent dimensions. Typically, Σ_ρ is parameterized using a scalar correlation parameter ρ , so that:

$$\Sigma_\rho = (1 - \rho)I_K + \rho \mathbf{1}_K \mathbf{1}_K^T.$$

This formulation allows each item's latent coordinates to be correlated with one another, which can capture global item traits shared across all K dimensions.

Item-Specific Shift α_j

For each item $j \in M$, we introduce a scalar intercept or preference parameter $\alpha_j \in \mathbb{R}$ that shifts the transformed latent coordinates equally. The transformation from the latent variable $U_{j,:}$ to utility vector $\eta_{j,:} \in \mathbb{R}^K$ is given by:

$$\eta_{j,:} = G^{-1}(\Phi(U_{j,:})) + \alpha_j \mathbf{1}_K^T,$$

where:

- $\Phi(\cdot)$ is the standard normal CDF, applied elementwise,
- $G^{-1}(\cdot)$ is an inverse link or transformation function (e.g., quantile function),
- α_j is a scalar item-specific covariate.

Prior on K

We often assume that the dimension K follows a truncated Poisson distribution with rate parameter λ , truncated to exclude zero (i.e., $K \geq 1$). The probability mass function is:

$$\pi(K = k) = \frac{e^{-\lambda} \lambda^k / k!}{1 - e^{-\lambda}}, \quad \text{for } k = 1, 2, \dots$$

This prior places most of the probability mass on small or moderate values of k , while still allowing some mass on larger dimensions, depending on the value of λ .

The posterior distribution

The general posterior is complicated than that, we are updating:

$$\pi_S(\rho, \beta, K, U \mid Y) \propto \pi_R(\rho) \pi_B(\beta) \pi_K(K) \pi(U \mid \rho, K) p_S(Y \mid h(U, \beta))$$

3.2 Reversible Jump-MCMC

Up-Move ($K \rightarrow K + 1$):

- Randomly select a column position from $\{0, 1, \dots, K\}$ at which to insert a new column.
- Generate a new “bridging column” $U_{:, \text{new}} \in \mathbb{R}^n$, typically from a prior distribution.
- Form the proposed matrix $U' \in \mathbb{R}^{n \times (K+1)}$ by inserting $U_{:, \text{new}}$ at the chosen position. The forward transition probability distribution is given by

$$Q(U'_0, \dots, U'_A, k+1 \mid U_0, \dots, U_A, k) = \rho_{k+1, k} \cdot \pi(U_{:, i} \mid U) \cdot \frac{1}{k+1}.$$

- Then accounting for the reverse probability of transforming from a $(k+1)$ -dimensional state back to a k -dimensional state. Pick a component at random from $k+1$ and delete it. The probability to propose this exact reverse move is

$$Q(U_0, \dots, U_A, k \mid U'_0, \dots, U'_A, k-1) = \rho_{k+1, k} \cdot \frac{1}{k+1}.$$

- Thus, the acceptance rate for the “up” move is

$$\alpha^+ = \min \left\{ 1, \frac{\pi(U', k+1) \times P(y \mid U', P) \times Q(U, k \mid (U', k+1))}{\pi(U, k) \times P(y \mid U, P) \times Q((U', k+1) \mid (U, k))} \right\}$$

- If the distribution on the extended space $\{U'\}$ is such that $\pi(U')$ is exactly the marginal density for the new vector, then from a statistical standpoint, we want: $\pi(U') = \pi(U, U_{:, i} = \pi(U) \pi(U_{:, i} \mid U)$.
- Accept or reject based on the Metropolis–Hastings acceptance ratio:

$$\alpha = \min \left\{ 1, \frac{\pi(k+1) P(y \mid U', P)}{\pi(k) P(y \mid U, P)} \right\}.$$

Down-Move ($K \rightarrow K - 1$):

- Randomly select one of the K columns of U to remove.
- Form the proposed matrix $U' \in \mathbb{R}^{n \times (K-1)}$ by deleting the selected column. The forward transition probability for the “down” move is

$$Q((U', k-1) \mid (U, k)) = \rho_{k, k-1} \times \frac{1}{k} = \frac{1}{2} \times \frac{1}{k},$$

We pick “down” with probability $\frac{1}{2}$ and then choose one column among k with probability $\frac{1}{k}$.

- To reverse the dimension-down move, generate a new column from the current distribution $\pi(U'_{:, i} \mid U)$ and append it back by inserting one of the $k-1+1$ positions.
- Accept or reject based on the reverse Metropolis–Hastings ratio:

$$\alpha^- = \min \left\{ 1, \frac{\pi(K-1) p(y \mid U')}{\pi(K) p(y \mid U)} \right\}.$$

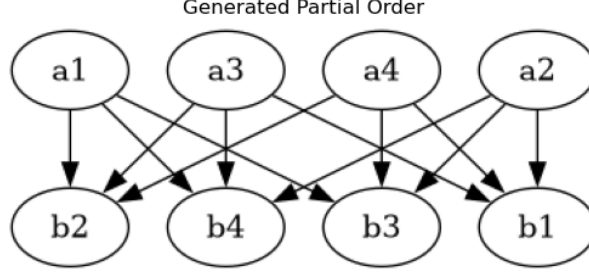


Figure 1: crown partial order with $n=8$

4 MCMC Examples: Crown Graph

In this section, we illustrate the Metropolis–Hastings approach with a crown poset, a classic example that achieves high dimension. Our goal is to show how the sampler behaves both with a fixed latent dimension K and with reversible-jumping MCMC.

4.1 Generating the Crown Poset S_{2k}

First, we construct a crown poset S_{2k} by partitioning $2k$ items into two disjoint sets:

$$A = \{a_1, a_2, \dots, a_k\}, \quad B = \{b_1, b_2, \dots, b_k\}.$$

Because of its highly symmetric and nonchain structure, the crown poset achieves the **maximum possible dimension** for a poset of size $2k$, namely $\dim(S_{2k}) = k$.

There are 16 Critical pairs with a realizer size of 4 for this crown graph.

After constructing the poset structure, we simulate 100 total orders with random length consistent with the partial order S_{2k} . This data serves as input for downstream inference tasks such as estimating the latent dimension K via MCMC.

4.2 Running the MCMC

For a crown poset with $n = 8$, the natural choice of latent dimension is $k = n/2 = 4$, which is known to be the exact dimension of the poset.

1. Fixed-Dimension MCMC Chains: We run separate MCMC chains with:

- $k = 2$ (underparameterized model),
- $k = n/2 = 4$ (true dimension).

For each chain, we compute the average inferred partial order matrix:

$$\hat{h}_{k,i,j} = \frac{1}{T} \sum_{t=1}^T h_{i,j}^{(k)}(t),$$

where $h_{i,j}^{(k)}(t)$ is the indicator that item $i \prec j$ in the inferred partial order at iteration t , for dimension k . We then compare the matrices across chains.

Expectation: If the fixed-dimension chains are recovering consistent structures, the corresponding $\hat{h}_{k,i,j}$ should lie approximately near the diagonal $y = x$, suggesting agreement on the relative ordering of items. The main outliers are the difference for edges between the two inferred partial orders.

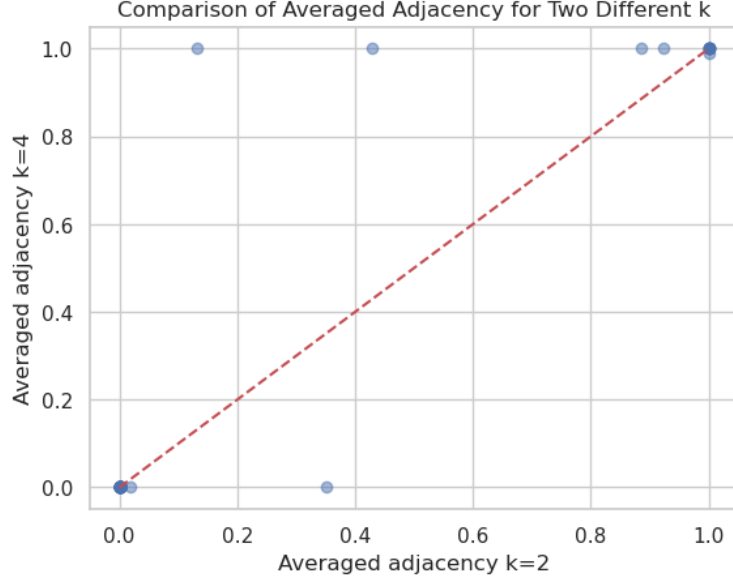


Figure 2: Averaged Adjacency for two different k

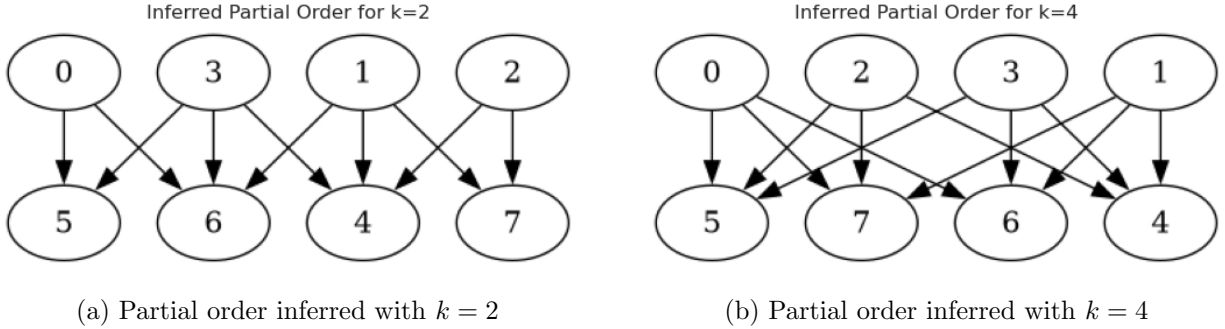


Figure 3: Comparison of inferred partial orders under different fixed dimensions.

2. Reversible Jump MCMC: We allow the dimension k to vary dynamically within the MCMC via reversible jump moves. This enables the model to explore and adapt the complexity of the latent space according to the observed data.

In our experiments, the RJMCMC procedure successfully recovered the true underlying structure of the crown poset. The posterior distribution over k exhibited a clear mode at $k = 4$, which matches the known true dimension of the crown.

5 Further Implications

Dimension K is *the* core parameter linking classical poset theory with modern Bayesian modeling. Although computing the exact dimension is NP-hard in general, Reversible Jump MCMC (RJMCMC) provides a practical framework for placing a prior over K and exploring the dimension space via stochastic proposals.

This exploratory approach enables us to assess how strongly the data support simpler versus more complex partial orders, balancing model flexibility with parsimony in a fully Bayesian manner. With more complicated k , we require more observations to reflect the all exist critical pairs.

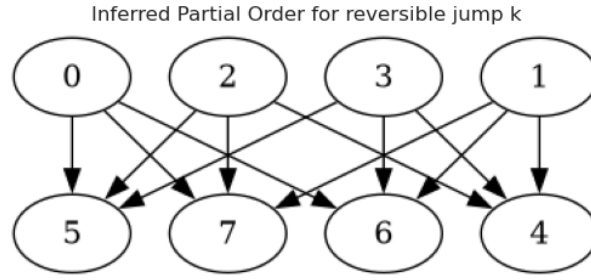


Figure 4: Partial Order Reversible Jump

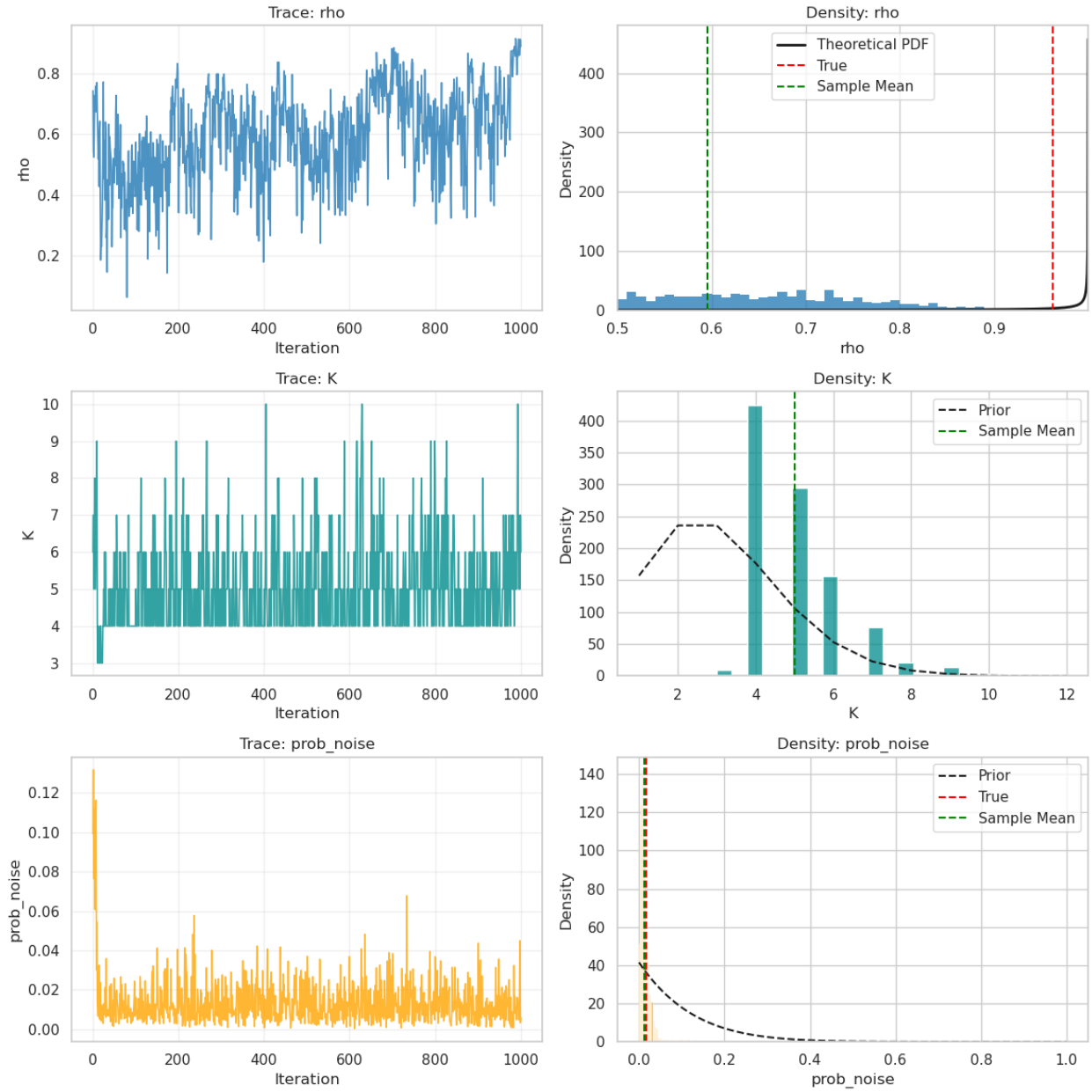


Figure 5: Posterior distribution of reversible MCMC

6 Appendix

6.1 The issue with the *no data* prior truncated k plot:

From a purely *prior* perspective—with **no data**—we would expect the long-run frequency of k to match the truncated Poisson($\lambda = 3$) distribution, shown by the red curve.

Indeed, the sampler’s acceptance/rejection mechanism *should* reflect those prior odds. However, there are two key RJ-MCMC details that systematically reduce the time spent at $k = 1$, making $k = 2$ and $k = 3$ relatively over-represented. The reasons are as follows: The truncated Poisson(λ) pmf is given by:

$$\pi_{\text{trunc}}(k) = \frac{e^{-\lambda} \lambda^k / k!}{1 - e^{-\lambda}}, \quad k = 1, 2, \dots$$

If λ is large (e.g. $\lambda = 3$), the prior at $k = 1$ is not overwhelmingly large. In fact, for $\lambda = 3$, the truncated Poisson probabilities for $k = 1, 2, 3, \dots$ are approximately:

$$k = 1 : 0.149, \quad k = 2 : 0.224, \quad k = 3 : 0.224, \quad k = 4 : 0.168, \text{ etc.}$$

- One-Way Boundary at $k=1$:

Going up from $k = 1$ to $k = 2$ has a high acceptance probability (if $\lambda = 3$, the acceptance ratio becomes “capped” at 1 in a Metropolis–Hastings accept step, effectively guaranteeing acceptance). In contrast, moving down from $k = 2$ to $k = 1$ has a smaller acceptance ratio:

$$\frac{\pi(1)}{\pi(2)} \approx 0.67 \quad (\text{for truncated Pois}(3)),$$

and only gets proposed half the time for the splitting “up vs. down” 50/50.

- Proposal Mechanics Tend to Favor Larger k

Even with no likelihood contribution (null data), each dimension-jump is a random-walk proposal:

At $k \geq 2$: Because $\lambda = 3$ is relatively large, “up” moves from $k = 2$ or $k = 3$ often have decent acceptance (e.g., from $k = 2$ to $k = 3$, ratio ≈ 1 ; from $k = 3$ to $k = 4$, ratio $3/4 = 0.75$).

At $k = 1$: Always “up” with an acceptance ratio usually ≥ 1 .

Hence, the chain tends to walk into the $k = 2, 3, 4$ zone, and though it can still move back down, the boundary effect at $k = 1$ plus half-time proposals for “down” makes smaller k states somewhat less visited.

And this is what the plot shows with no data plot shows: That “excess” around 2–3 is a common artifact of RJMCMC when there is a hard limit at $k = 1$. Numerically, the chain sample mean (≈ 3.34) is still quite close to the theoretical 3.16, but the dynamics of the random walk proposal produces a small upward bias away from $k = 1$. This does not necessarily mean your chain is “wrong”; it’s just a well-known quirk of dimension-jumping MCMC with a boundary at 1.

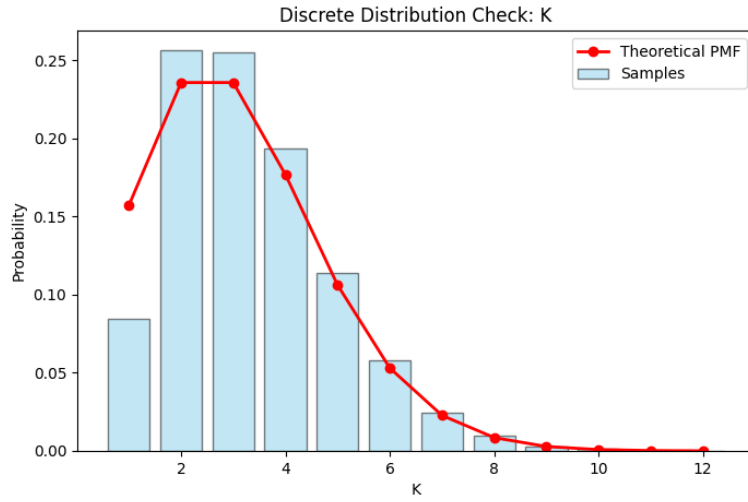


Figure 6: The RJ MCMC K plot with no data

References

- Ben Dushnik and E. W. Miller. Partially Ordered Sets. 63(3):600. ISSN 00029327. doi: 10.2307/2371374. URL <https://www.jstor.org/stable/2371374?origin=crossref>.
- Mihalis Yannakakis. The Complexity of the Partial Order Dimension Problem. 3(3):351–358. ISSN 0196-5212, 2168-345X. doi: 10.1137/0603036. URL <https://epubs.siam.org/doi/10.1137/0603036>.