Metropolis-Hastings Algorithm: Exploring Computational Poetry Through Simulation

Saeed Mohseni seh deh
The Bradley Department of Electrical and Computer Engineering at Virginia Tech

1 Introduction

In this brief simulation study, the Metropolis-Hastings algorithm, one of the most significant algorithms of the 20th century, is investigated. The objective of this algorithm is to enable the drawing of samples according to a probability density function (PDF) for which the equation is only available through samples drawn from a PDF from which we can sample. It is a Markov Chain Monte Carlo (MCMC) algorithm and has applications in various domains such as Bayesian Inference, Statistical Physics, Machine Learning, and others. In this simulation study, this algorithm is utilized to create samples of the t-distribution with ν degrees of freedom as the target distribution from a normal PDF as the proposed distribution, and vice versa. Numerous scenarios in terms of the number of samples and different distribution parameters are considered for each case. Additionally, the conditional proposal form for sampling from the t-distribution using the normal PDF is also examined. This report is structured as follows: Section 2 is dedicated to the simulation setup, results, and discussion, while Section 3 concludes this report.

2 Simulations

This section is devoted to the presentation of simulation results. Simulations can be categorized into three groups:

1. Sampling from the target distribution of t with ν degrees of freedom by utilizing the proposed Gaussian distribution with a mean of zero and a variance of one.
2. Sampling from the target distribution of t with ν degrees of freedom by utilizing the proposed conditional Gaussian distribution with the current accepted sample as the mean and a variance of one.
3. Sampling from the target distribution of a Gaussian with a mean of zero and a variance of one by utilizing the proposed distribution of t with ν degrees of freedom.

In each of the simulations, the number of resulting samples from the algorithm and degrees of freedom are varied. Specifically, $N \in 10, 100, 1000, 10000$ and $\nu \in 2, 5, 10, 20$, where N and ν denote the number of samples and degrees of freedom, respectively. For the reporting purpose, the Histogram and Quantile-Quantile plots (QQ plots) are utilized, which are useful visual tools for comparing the similarity of two distributions. Figure 1 and 2 demonstrate the simulation results for the first group, Figs 3 and 4 display the simulation results of the second group, and Figs 5 and 6 illustrate the simulation results of the third group.

2.1 Discussion

In this part, a brief discussion of the simulation results is presented. Figures 1-4 show the result when the t-distribution is created using samples from the normal PDF. As can be observed, when the number of samples, i.e., the number of iterations of the algorithm, increases, the results become more accurate. This can be seen from both the histogram plots and QQ plots. The primary reason is that when the number of samples increases, those samples will cover a broader range of the PDF. To elaborate, the probability of obtaining a very large or very small value is minimal for the normal distribution, and if no samples come from those regions, it is not possible to create the tails of the t-distribution accurately. However, if the number of drawn samples increases, the likelihood of obtaining values in the tails of the normal PDF also increases, and the behavior of the tails of the t-distribution is captured more precisely.

Another noteworthy observation in creating the samples of the t-distribution from normal samples is that as the degrees of freedom of the target t-distribution increase, the algorithm performs better. That is, as the degrees of freedom of the t-distribution grow, it becomes more similar to the normal PDF, thus the target and proposal distributions become more alike, and the performance of the algorithm improves.

Another point can be revealed by comparing Figs 1, 2 with Figs 3, 4. The first set of figures demonstrates the results of the algorithm when an unconditional normal PDF with a mean of zero and a variance of one is used, and the second set of figures shows the result when a conditional normal PDF with the mean set to the latest accepted sample and a variance of one is used. It can be observed that the conditional version performed better. The reason for this can be explained as follows. It is known that the t-distribution has heavier tails than the normal distribution, and it is in these tails where these distributions differ significantly. Imagine a sample in the tail of the normal distribution is produced and accepted by the algorithm. This sample can significantly improve the information about the target distribution. By centering the proposed distribution there, it becomes more likely to draw a sample in the tail, as the mean is in the tails. Thus, this technique facilitates faster convergence of the algorithm and, in a limited number of iterations, enables it to achieve better results.

Figures 5 and 6 show the result of the last group of simulations, creating a normal distribution from the t-distribution. As the figures demonstrate, the performance of the algorithm in this case is much better. The reason simply lies in the fact that the t-distribution has heavier tails and better coverage of samples resulting from sampling that distribution. This high coverage of samples captures all the behavior of the normal PDF more accurately.

3 Conclusion

This simulation study explored the Metropolis-Hastings algorithm for sampling from t-distributions and normal distributions. The results demonstrated the algorithm's effectiveness in accurately approximating these target probability density functions. Increasing iterations and degrees of freedom improved performance when sampling t-distributions using a normal proposal, due to better coverage of the target range. Utilizing a conditional normal proposal centered on accepted samples further enhanced convergence, especially for heavy-tailed targets. Sampling a normal from a t-distribution exhibited excellent performance, leveraging the heavier tails of the proposal for comprehensive target coverage with fewer iterations. Overall, this study highlights the versatility of the Metropolis-Hastings algorithm and reaffirms its importance as a fundamental sampling tool across computational statistics and related fields.

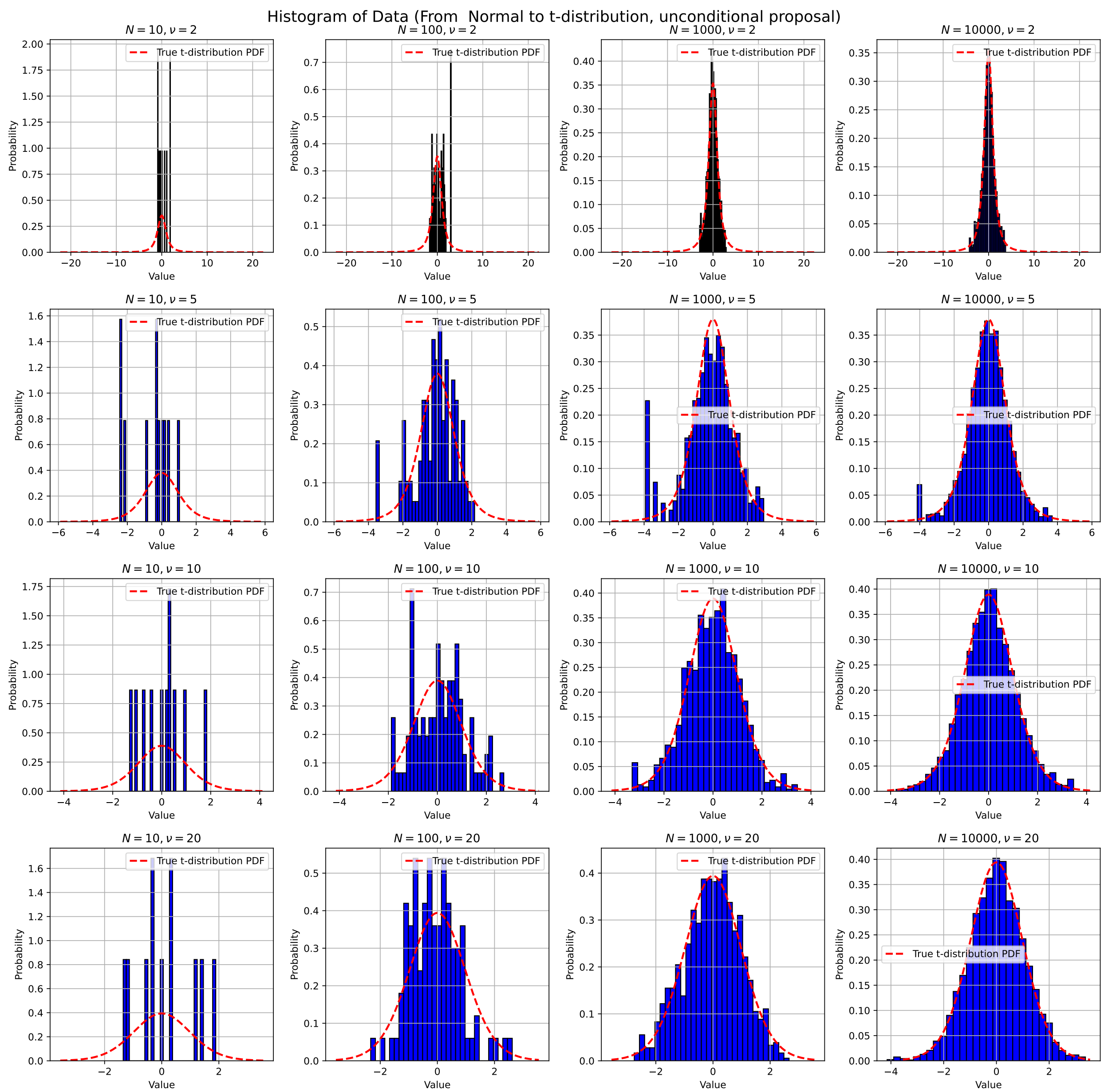

Figure 1: Target distribution: t with ν degrees of freedom, proposed distribution: normal with mean 0 and variance 1

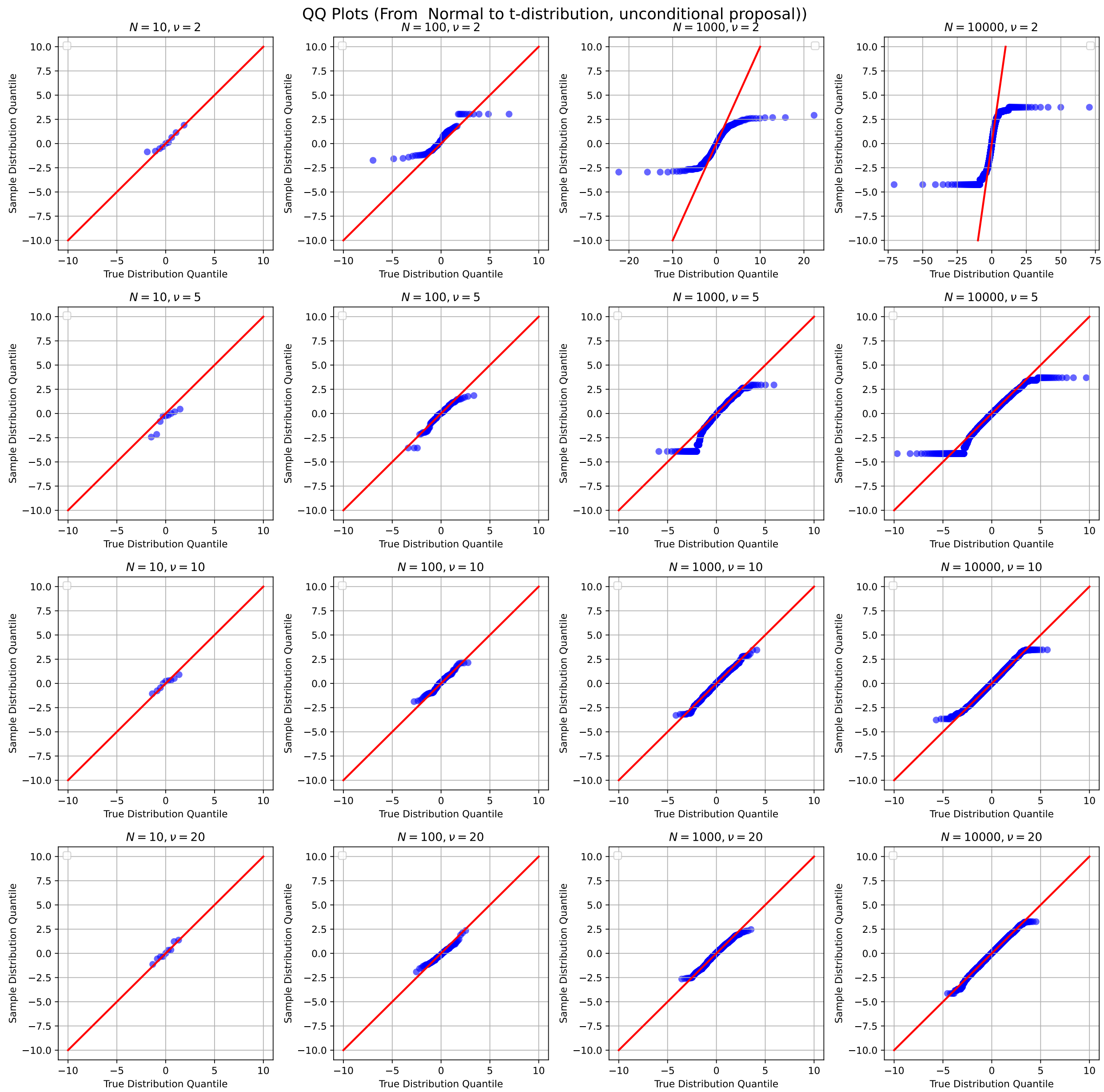

Figure 2: Target distribution: t with ν degrees of freedom, proposed distribution: normal with mean 0 and variance 1

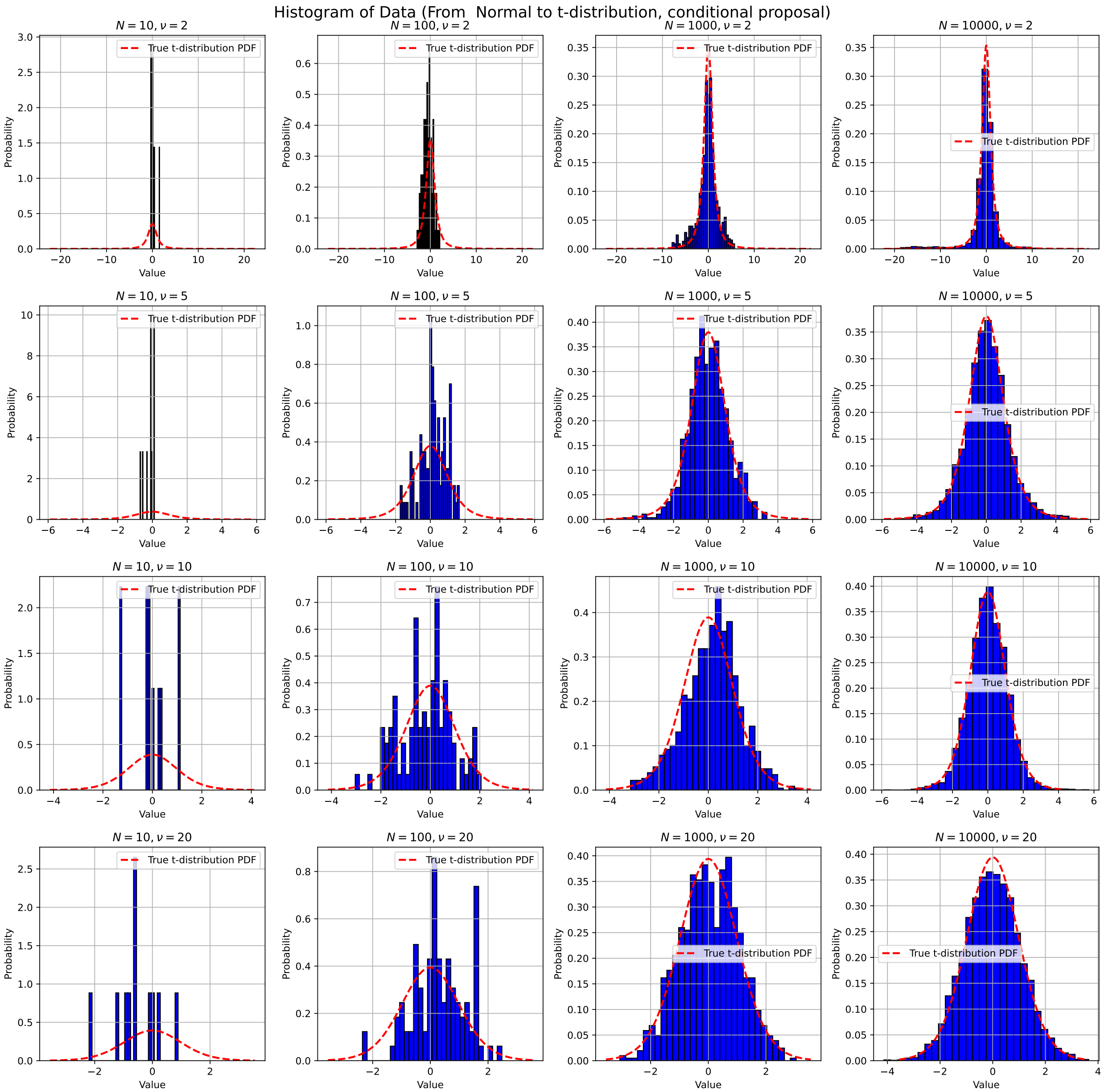

Figure 3: Target distribution: t with ν degrees of freedom, proposed normal with current sample as mean and variance 1

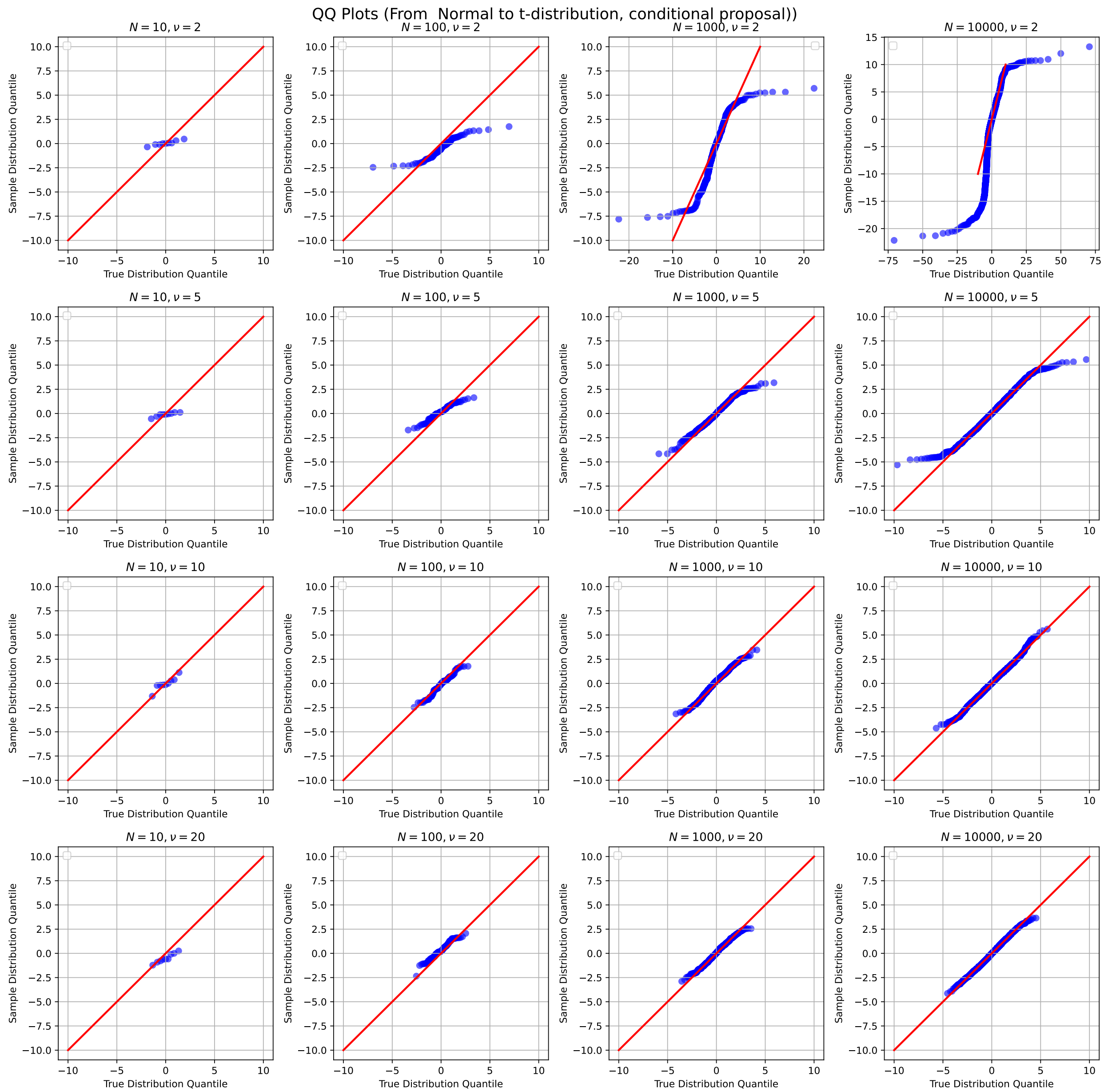

Figure 4: Target distribution: t with ν degrees of freedom, proposed distribution: normal with current sample as mean and variance 1

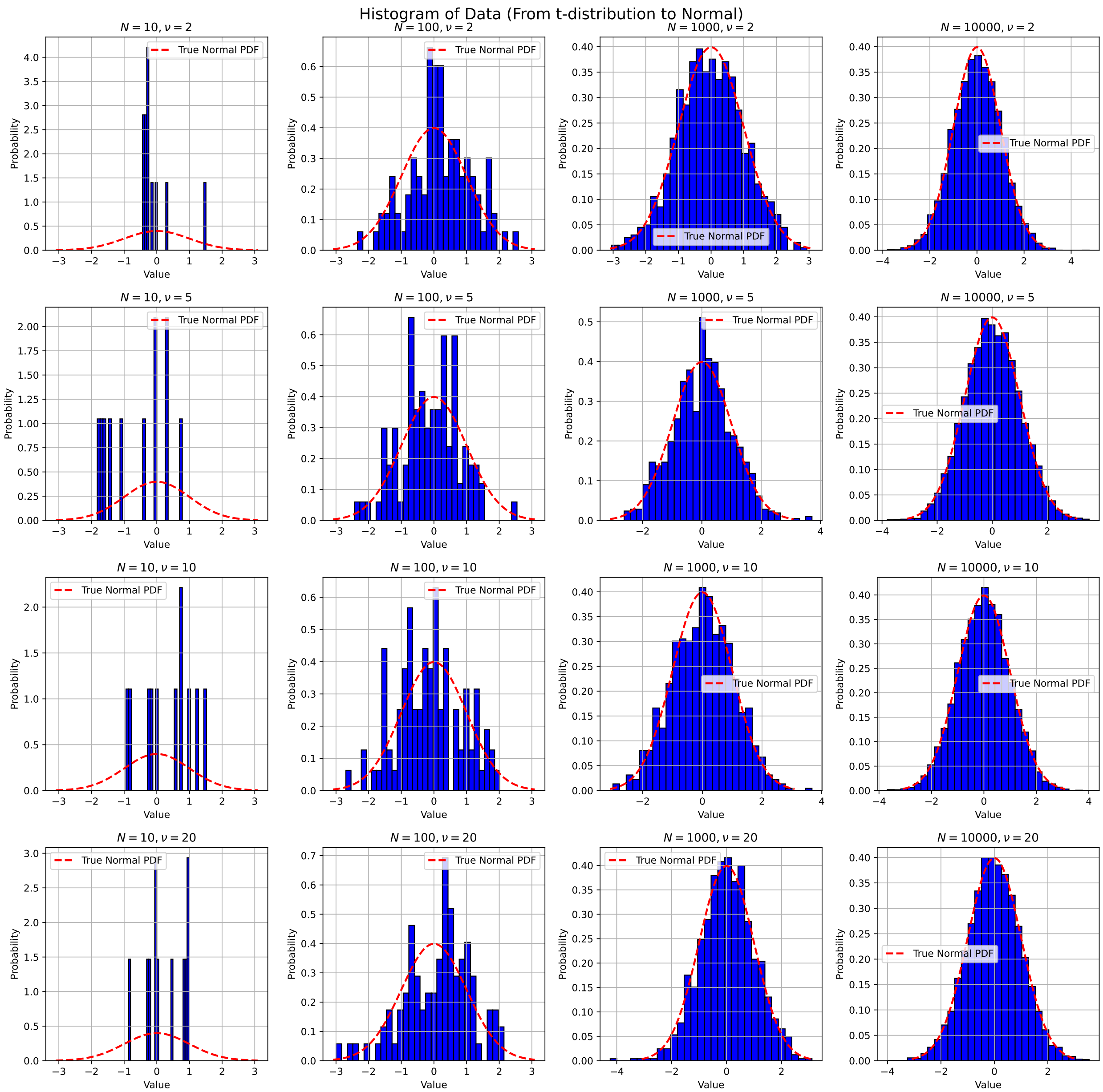

Figure 5: Target distribution: normal with mean 0 and variance 1, proposed distribution: t with ν degrees of freedom

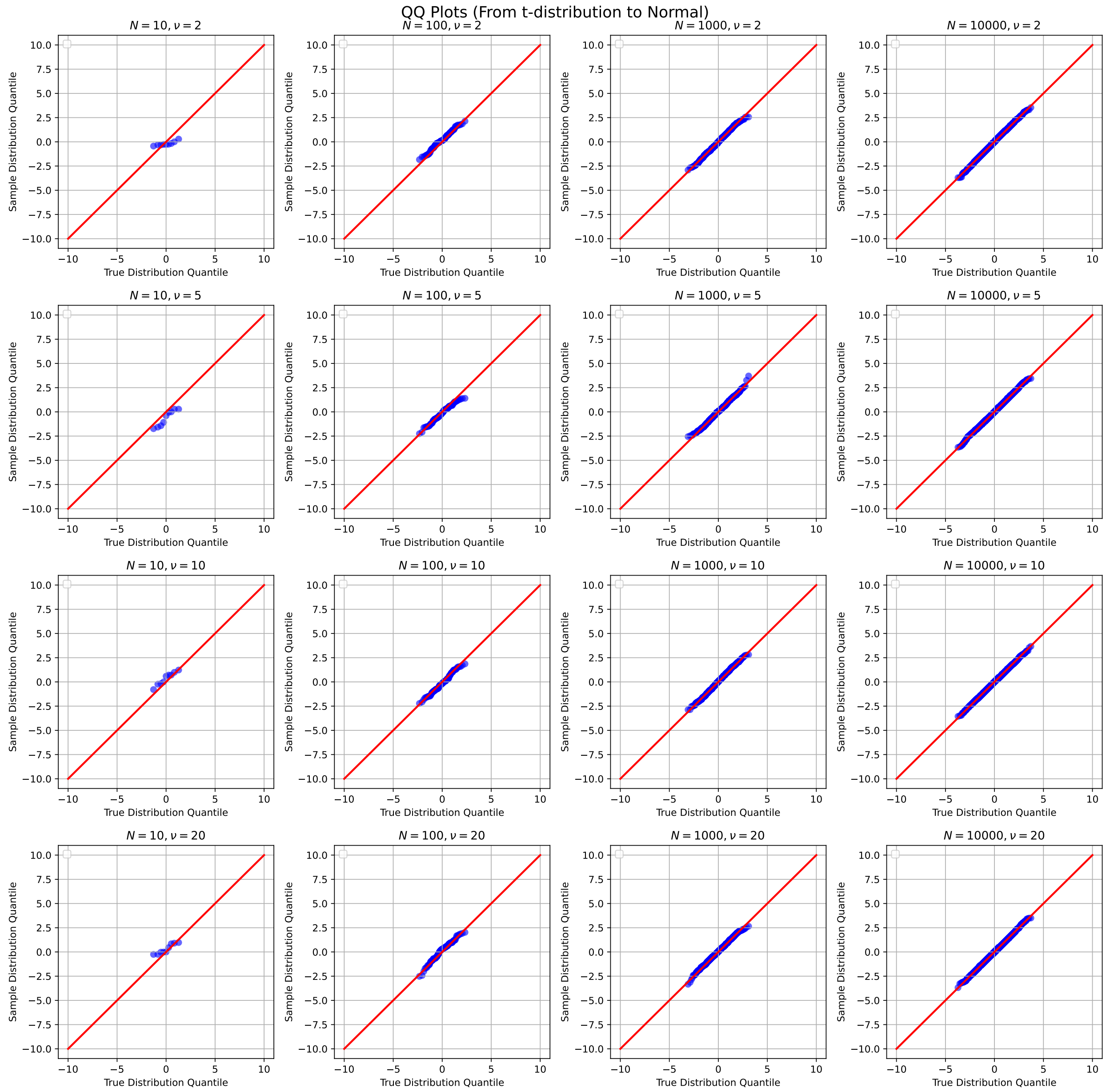

Figure 6: Target distribution: normal with mean 0 and variance 1, proposed distribution: t with ν degrees of freedom