THE CHOICE OF AUXILIARY DENSITY FUNCTION IN STOCHASTIC COLLOCATION

KELLIN RUMSEY & NEVIN MARTIN

12/06/2016

1. Introduction

Stochastic collocation has been used in a wide range of application spaces as a way to efficiently propagate the uncertainty in a set of input parameters through a mathematical model. There are certain difficulties that arise when the random variables involved are dependent on eachother. A solution proposed by [1] involves using auxillary random variables which are independent, and obtaining the correct solution based on these. In this paper, we will use a simple illustrative example to examine the choice of these independent random variables, and their effect on the computational cost of Stochastic Collocation.

1.1. **Stochastic Collocation.** In stochastic collocation, we have

$$E[u(\bar{x}, Y)] = \int_\Gamma u(\bar{x}, Y)\pi(Y)dY \approx \sum_{m=1}^{\eta} w_m U_h(\bar{x}, Y^{(m)}) \tag{1}$$

where

$$w_m = \prod_{n=1}^{N} \int_{\Gamma_n} L_{m_n}(y_n)\pi_n(y_n)dy_n \tag{2}$$

Here, w_m and $y^{(m)}$ are the weights and collocation points, respectively, that are determined using a joint probability density function (pdf) of N random variables. If the random variables are independent, it is generally straightforward to find these values. However, if the random variables are dependent, we have $\pi(y) \neq \prod_{n=1}^{N} \pi_n(y_n)$. This can make it very difficult to evaluate the integral shown in Equation 2.

1.2. **Auxiliary PDFs.** To combat this issue, Babuška, Nobile, and Tempone (2007) proposed to use an auxiliary pdf $\hat{\pi} : \Gamma \to \Re_+$ as the joint pdf of N random variables, as shown in Equation 3.

$$\hat{\pi}(y) = \prod_{n=1}^{N} \hat{\pi}_n(y_n) \tag{3}$$

Using this auxiliary pdf, the expected value can be approximated using the summation below:

$$E[u(\bar{x}, Y)] = \int_\Gamma u(\bar{x}, Y)\pi(Y)dY = \int_\Gamma u(\bar{x}, Y)\frac{\pi(\hat{Y})\pi(Y)}{\pi(\hat{Y})}dY \approx \sum_{m=1}^{\eta} w_m \frac{\pi(Y^{(m)})}{\hat{\pi}(Y^{(m)})} U_h(\bar{x}, Y^{(m)}) \tag{4}$$

where the weights w_m and collocation points $y^{(m)}$ are now found using $\hat{\pi}(y)$ in equation (2).

1.3. **An Illustrative Example.** For this study, a simple ODE was used to illustrate the ability of an auxiliary pdf to approximate a pdf of dependent random variables. For $t \in [0, 1]$,

$$u_t = au \qquad u_0 = 1 \tag{5}$$

Of course, the solution here is

$$u(t) = e^{at} \tag{6}$$

In this ODE, the coefficient $a = a(y_1, y_2)$ was chosen to be a bounded function of two dependent random variables, such that:

$$a(y_1, y_2) = 2 - e^{\frac{-1}{2}(y_1^2+y_2^2)} \tag{7}$$

with

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \sim BVN(\mathbf{0}, \Sigma) \tag{8}$$

where Σ is the covariance matrix, as shown below:

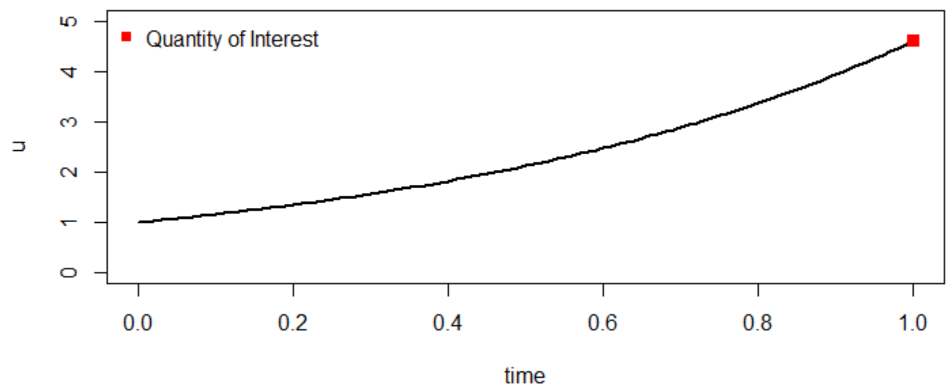
$$\Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \tag{9}$$

Where $\rho \in (0, 1)$ represents the correlation of y_1 and y_2, and was varied throughout the study. Additionally, it was decided that $Q(y_1, y_2) = \mathbb{E}[u(t = 1)]$ would be the Quantity of Interest (QoI) to be studied, as shown in Equation 10.

$$Q(Y) = \int_{\mathbb{R}} \int_{\mathbb{R}} e^{a(y_1,y_2)} \pi(Y) dy_1 dy_2 \tag{10}$$

An accurate solution to this QoI was obtained using Hermite Quadrature with many collocation points, the result of which can be seen in Figure 1. This is used as a reference solution throughout the remainder of the analysis.

FIGURE 1. A Highly Accurate Reference Solution



2. CONSIDERATIONS

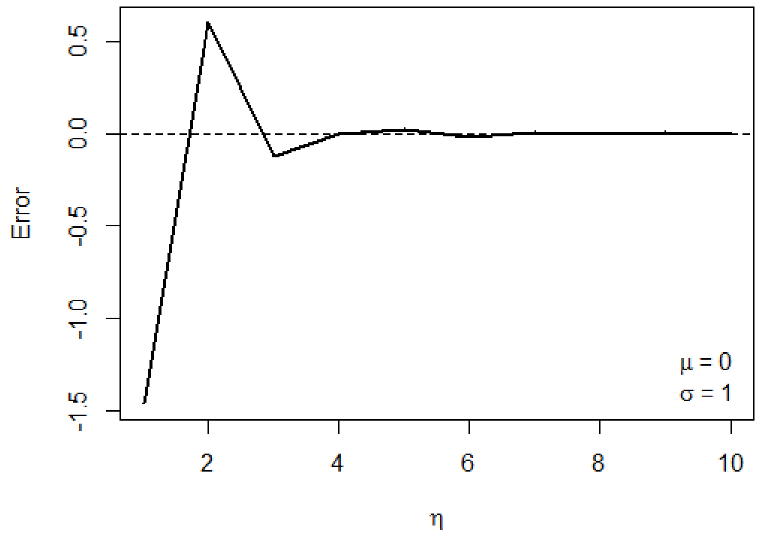
For this study, we decided on three important considerations when choosing an auxillary pdf. We consider the effects of support and shape of $\hat{\pi}(Y)$ on the solution, and we also consider the strength of correlation between the variables in $\pi(Y)$. In order to closely examine these properties, Normal, Uniform and Beta pdfs were used as auxillary pdfs in our analysis.

2.1. **Support.** For the first two considerations, we fix $\rho = 0.5$. Since the random variables in a multivariate normal are themselves marginally Normal, the natural choice for an auxillary pdf here is $N(0, 1)$. Two independent $N(0, 1)$ variables were used to construct the auxiliary density function. This also provides a good base case for this section, since the stochastic space Γ is the same for π and $\hat{\pi}$.

Figure 2 shows the error as a function of collocation points for the N(0,1) pdf. From this plot, it can be seen that the error quickly converges as the number of collocation points increases. Only 10 collocation points are required for convergence when $\epsilon_{TOL} = 0.001$. Indeed, the $N(0, 1)$ density appears to be an appropriate function to use for the auxiliary pdf in this case.

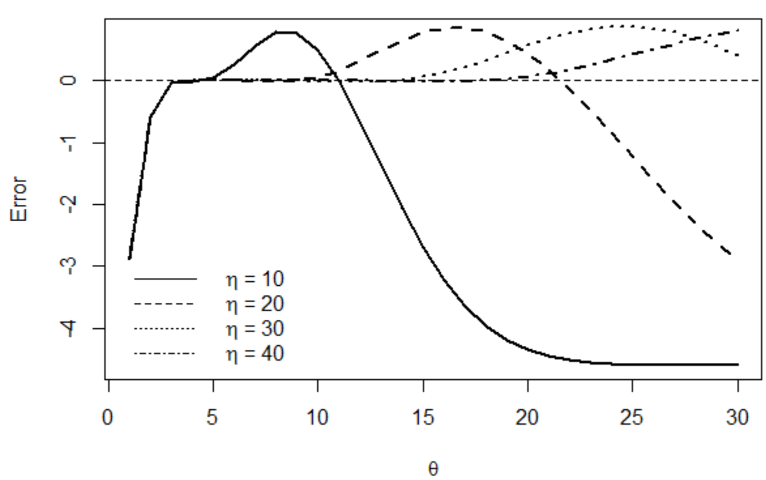
To study the effects of support on the auxillary solution, we will use independent random variables of the form $U(-\theta, \theta)$. Figure 3 shows the error of the approximated solution against θ for different numbers of collocation points (η). For very small values of θ, we are unable to obtain an accurate solution, regardless

FIGURE 2. Number of Collocation Points Required for Convergence of $N(0,1)$



of the number of collocation points. As θ increases to approximately 5, there appears to be convergence for as little as 20 collocation points. However, as θ increases, the number of collocation points needed for convergence also increases.

FIGURE 3. Number of Collocation Points Required for Convergence of $U(-\theta,\theta)$

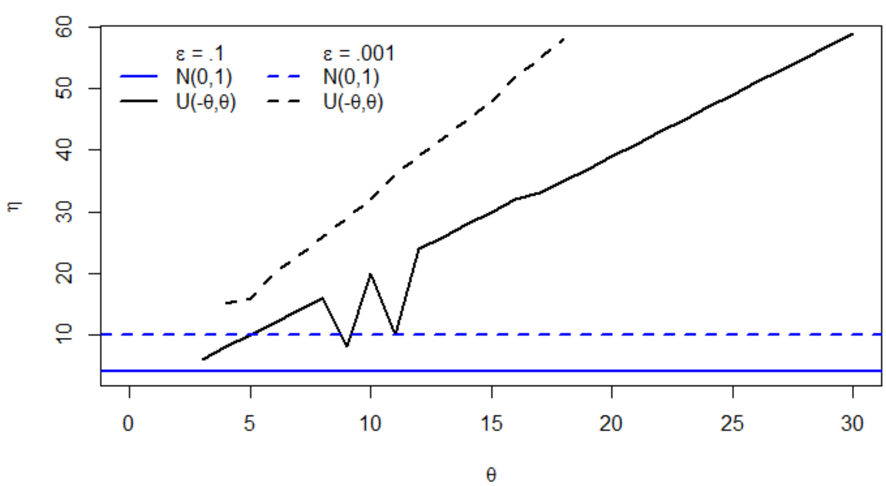


In fact, the relationship between θ and η is linear. As Figure 4 demonstrates, the number of collocation points needed for convergence is proportional to θ. We also see that the slope of this linear relationship is also dependent on ϵ_{TOL}, a thought that will be discussed more in the final section of this paper.

A result which is less evident from Figure 4, is that the value θ required for convergence also increases with ϵ_{TOL}. For example, we can obtain a solution which is accurate within $\epsilon_{TOL} = 0.1$ using $\theta \approx 3$, but if we would like to be accurate within $\epsilon_{TOL} = 10^{-3}$, then we need $\theta \approx 4$. In other words, the number of collocation points η required to converge to ϵ_{TOL} for a given θ is given by equation (11). Where both c_ϵ and k_ϵ increase as $\epsilon_{TOL} \to 0$.

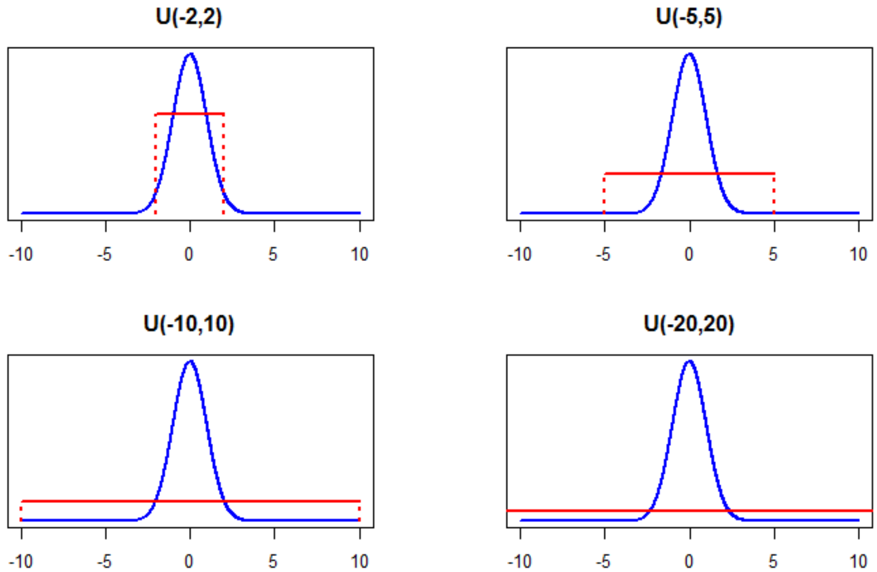
$$\eta = \begin{cases} \infty & \theta < c_\epsilon \\ k_\epsilon \theta & \theta \geq c_\epsilon \end{cases} \tag{11}$$

FIGURE 4. A Linear Relationship Between θ and η



We use Figure 5 to summarize our results. The top left image will not be able to converge to a very accurate solution at all, while the bottom right figure is capable of converging to a very precise solution. There is a trade-off of course; the bottom right figure requires many more collocation points in order to reach convergence. If independent Uniform random variables are to be used for $\hat{\pi}$, then we should choose an ϵ_{TOL}, and minimize θ so that the solution still converges.

FIGURE 5. Comparing $N(0,1)$ and $U(-\theta,\theta)$



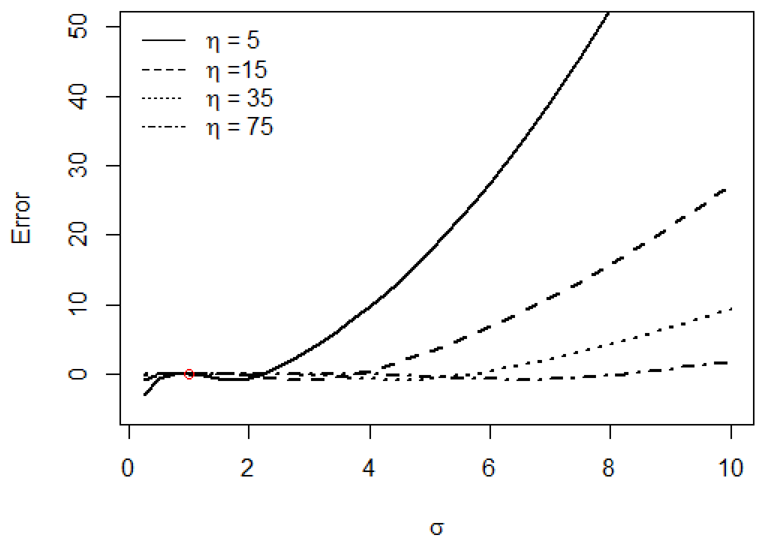
2.2. **Shape.** For the shape analysis, we fix $\rho = 0.5$ as before. Additionally, from the support analysis, it was found that a θ value of 5 gave efficient solutions for $\epsilon_{TOL} = 10^{-3}$; therefore, the analysis in this section uses a fixed value of $\theta = 5$.

We first return to the $N(0,1)$ case, and the effects of auxillary density shape was examined by changing the standard deviation. Figure 6 shows the error produced using Normal(0,σ) random variables for the auxiliary pdf for increasing levels of σ. As expected, we see that N(0,1) has the best convergence, with only 10 collocation points needed for $\epsilon_{TOL} = 0.001$. However, as the value of σ increases, as the does the number of collocation points required for convergence. At N(0,5) and the same level of error tolerance, approximately 35 collocation points are required.

Now we return to the case of bounded support. By recalling that the Uniform density is a special case of the Beta density, we utilize the Beta(α, β) distribution to see the effects of changing its shape parameters. Since the Beta distribution is normally restricted to the interval [0,1], we use a more general version which allows us to define support on the interval $[-\theta, \theta]$ ($\theta = 5$ for this section). It was also decided that the values of α and β should be equal to produce a symmetric distribution. Using a simple moment matching technique, we can determine a value for α which well approximates the Normal Density[1].

[1]A better approximation can be obtained by matching the peaks of the distribution, but this must be obtained by solving $\theta = \frac{\sqrt{2\pi}\sigma}{2\beta(\alpha,\alpha)}$ numerically.

FIGURE 6. Error vs σ for Different Values of η



$$\alpha = \frac{(\theta/\sigma)^2 - 1}{2} \tag{12}$$

For our purposes, we introduce the following notation for the Beta distribution: $B(\theta, \alpha)$, which is a $Beta(\alpha, \alpha)$ mapped to the interval $[-\theta, \theta]$. Figure 7 shows the error of a B(5,α) distribution for increasing values of α. Note that when $\alpha = 1$, the Beta distribution is the same as the Uniform distribution. When η = 5 (solid line), the Uniform ($\alpha = 1$) clearly has not converged. However, if the shape is changed to α = 50, there is reasonable convergence. Specifically, when $\epsilon_{TOL} = 0.001$, only 10 collocation points are needed when $\alpha = 50$, which is the same number of collocation points required by the N(0,1) pdf.

Figure 8 gives another visual representation of the intuition behind these results. With the Uniform distribution (upper left plot), the convergence is sub-optimal due to the fact that the auxiliary pdf is not accounting for the shape of the original distribution. As α and β increase, the auxiliary distribution is able to more accurately cover the same probability space as the original distribution, and with that, less collocation points are needed for convergence. In the case of ρ = .5 and θ = 5, an α parameter of 12 provides the optimal results, where $\alpha = 12$ is chosen according to equation (12).

2.3. **Correlation.** Finally, the effects of the correlation between the dependent random variables in the multivariate normal distribution were studied. Again, the value for θ was fixed to 5, while the value of ρ

FIGURE 7. Error vs α for $\theta = 5$ and Different Values of η

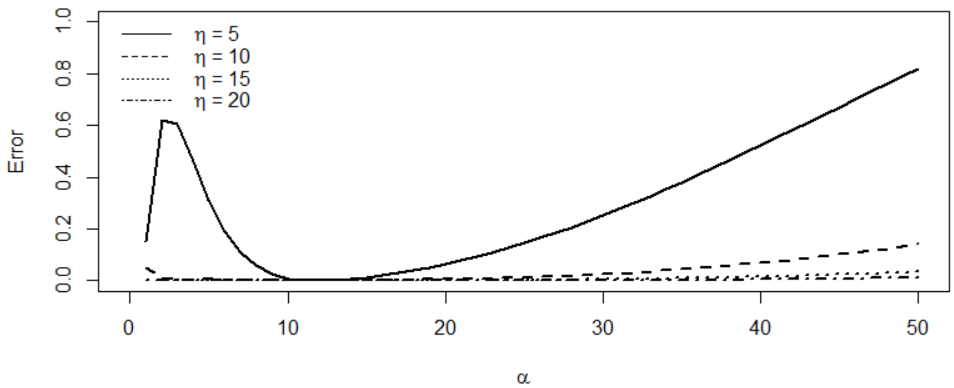
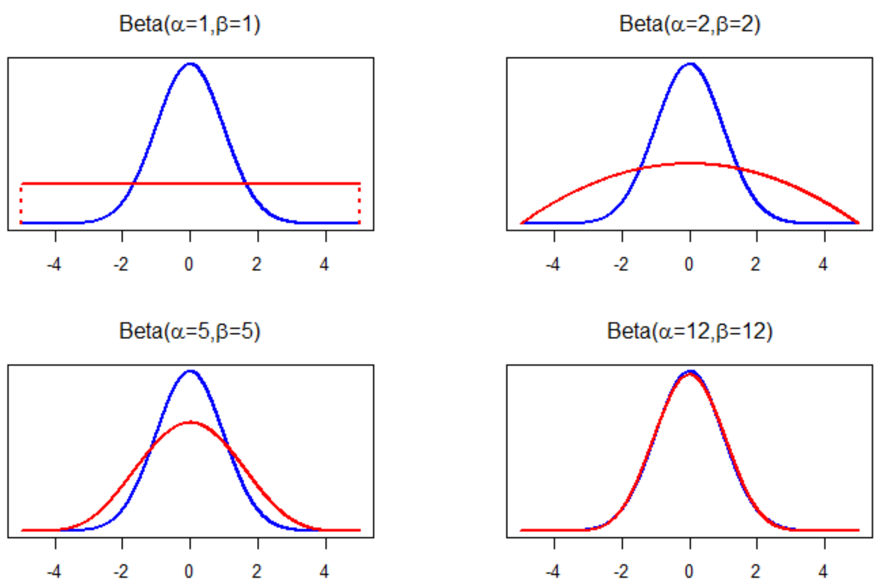


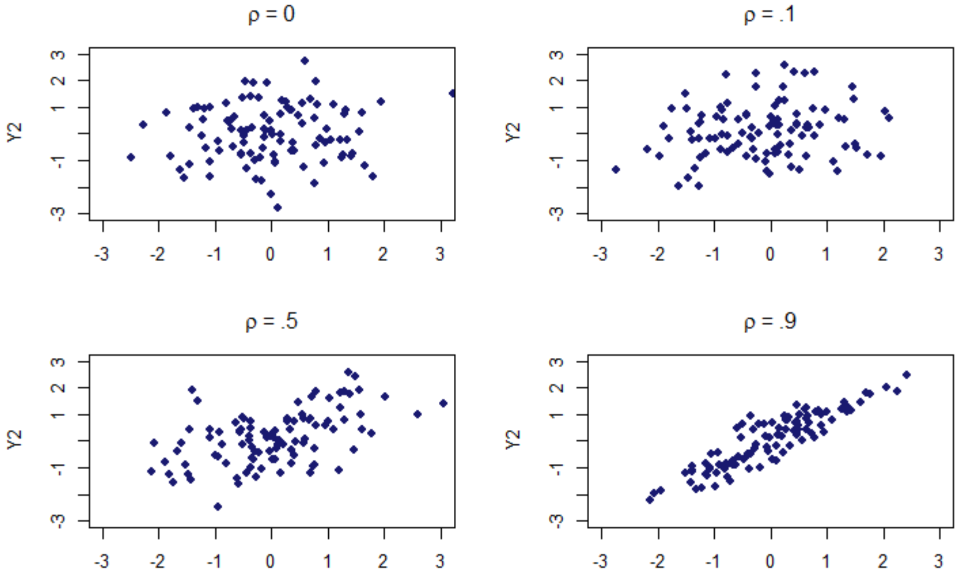
FIGURE 8. Comparing $N(0,1)$ and $Beta(\alpha, \alpha)$



varied from 0 to .9. Figure 9 shows the scatterplots of two Normal random variables (Y_1 and Y_2) with varying degrees of correlation. From the plot, it can be seen that the the most drastic change in the distribution occurs between $\rho = .5$ and $\rho = .9$.

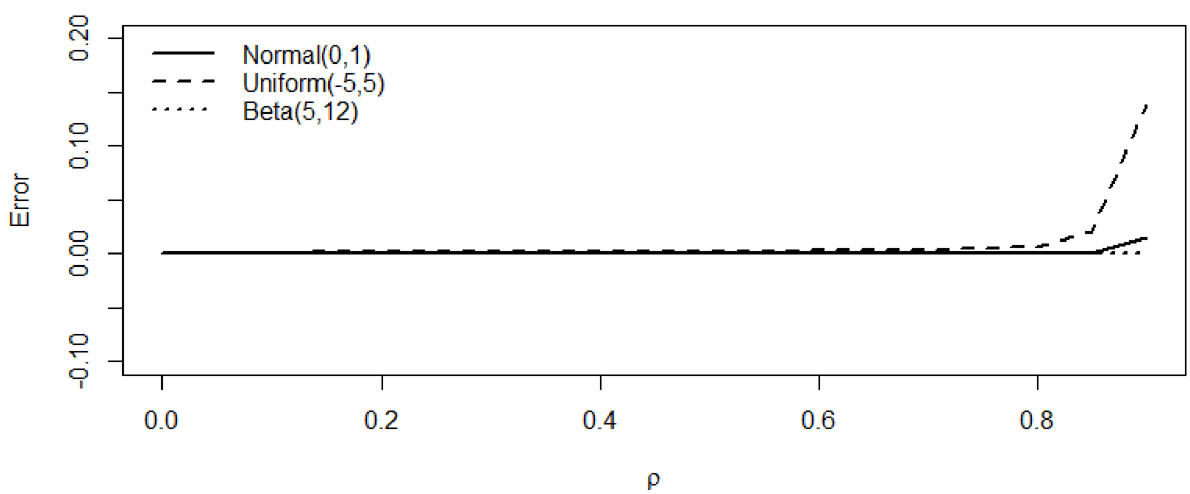
To see how the changes in the distribution due to varying ρ values affected the convergence of the auxiliary pdfs, the error of the three distributions was examined. Figure 10 shows the error of the Normal(0,1), Uniform(-5,5) and Beta(5,24) distributions using 15 collocation points. This figure supports what was seen in Figure 9; that is, for high levels of correlation, more collocation points are generally needed to accurately approximate the correlated data. Furthermore, this plot shows that while the uniform and normal

Figure 9. Bivariate Normal Samples for Different Values of ρ



distributions require more points when $\rho > 0.8$, the Beta distribution performs well at all correlation levels with only 15 collocation points.

Figure 10. Comparison of Error with $\eta = 15$



**Beta(5,12) corresponds to $\theta = 5$ and $\alpha = 12$)*

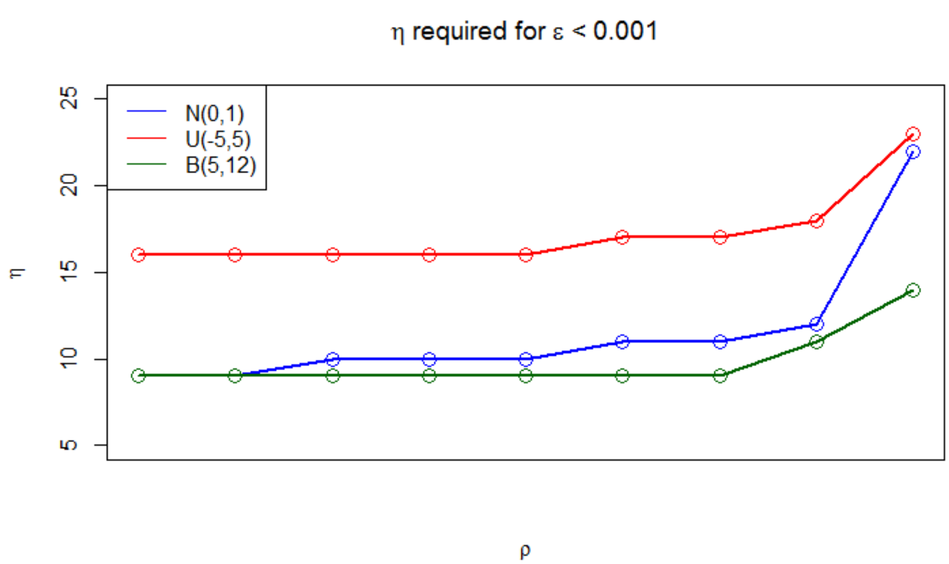
Figure 11 further illustrates how correlation effects the results convergence at $\epsilon_{TOL} = 0.001$. Most notably, the B(5,12) and Normal(0,1) pdfs require a similar number of collocation points for low and moderate levels of correlation. However, once ρ has reached 0.9, B(5,12) requires 8 less collocation points than the Normal distribution.

3. Results and Conclusion

We begin with a discussion of each of the three considerations. Using auxillary random variables which truncate the support, can be an efficient choice. However, we must be sure that the truncation parameter θ will give us convergence at our desired ϵ_{TOL}. Secondly, we showed that the shape of the auxillary density function has a significant impact on efficiency of Stochastic Collocation. Although, there is a caveat to this which will be addressed in this section. Finally, we see that in general, we will need more collocation points if the variables described by $\pi(Y)$ are highly correlated, and in this case we found that the Beta distribution can outperform the Normal Distribution.

Figure 11 provides a nice summary of these results in this simple illustrative example. Table 1 provides a similar summary.

Figure 11. Comparison of Cost for $\epsilon_{TOL} = 10^{-3}$ vs Correlation



Finally, we must return to some of the more delicate points of this discussion. Let us recall the relationship between θ and η when dealing with Uniform auxillary random variables.

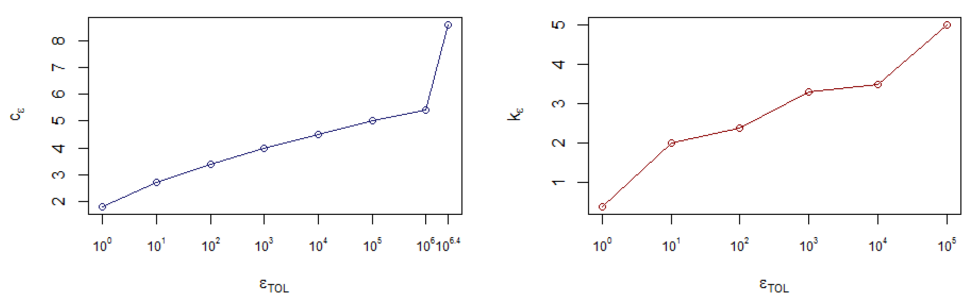
$$\eta = \begin{cases} \infty & \theta < c_\epsilon \\ k_\epsilon \theta & \theta \geq c_\epsilon \end{cases}$$

TABLE 1. Collocation Points Required to acheive $\epsilon_{TOL} = 10^{-3}$

ρ	0.1	0.5	0.9
N(0,1)	9	10	22
U(-5,5)	16	16	23
U(-10,10)	22	32	46
B(5,12)	9	9	14
B(10,49.5)	9	10	19

Analytic results for c_ϵ and k_ϵ may be obtainable, but for the sake of our illustration we provide the following plots, and table of values.

FIGURE 12. Emprical Values of k_ϵ and c_ϵ vs ϵ_{TOL}



We notice that the curve for c_ϵ appears to be leveling off, but rapidly shoots off to infinity. This could be due to one of two things. Either this method is unable to acheive precision smaller than approximately 10^{-6}, or the method is unstable in some way. More analysis is required to know for sure, but we see that good accuracy can be obtained for reasonably small θ, as long as $\epsilon_{TOL} > 10^{-6.3}$.

Suppose we wanted to use Independent Uniform auxillary random variables, and we desire an accuracy of $\epsilon_{TOL} = 10^{-4}$. Then the left panel of Figure 12 tells us that we need θ to be at least 4.5. The right panel of Figure 12 tells us that we shuld use at least $k_\epsilon \theta = (4.5)(3.5) = 15.75$ collocation points. Hence,the choice of $\theta = 4.5$ and $\eta = 16$ should provide optimal results. Here, we would recommend finding an optimal α, and applying the Beta shape correction for even better results. The values in Figure 12 are given in Table 2.

TABLE 2. Select Values of c_ϵ and k_ϵ

ϵ_{TOL}	10^0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	$10^{-6.4}$	10^{-7}
c_ϵ	1.8	2.7	3.4	4	4.5	5	5.4	8.6	∞
k_ϵ	0.4	2	2.4	3.3	3.5	5	-	-	-

We conclude with a short discussion of the shape improvement. We have discussed that using a Beta distribution rather than Uniform can cut the cost significantly. The caveat here, is that when θ is very close to the boundary, using a Beta distribution can perform far worse. When θ is just barely large enough to converge, this tells us that we need the information included in the tails of the normal distribution. The Beta shape correction may not give enough weight to this information in the tails. The solution here of course is to increase θ slightly and then apply the shape correction for best results. The cost of increasing θ is outweighed by the ability to efficiently apply the shape correction.

We believe there is room here for future research. Primarily these results should be generalized and supported with analysis. We believe that the approach of using independent Beta random variables is quite flexible and can be extended to practically any form for π, not just Multivariate Normal. We have shown that the support parameters and shape parameters can be chosen for optimality, and this method can easily be extended to the marginals of any multivariate density function.

References

Babuška, Ivo and Nobile, Fabio and Tempone, Raúl. (2007). *A Stochastic Collocation Method for Elliptic Partial Differential Equations with Random Input Data.* SIAM Journal on Numerical Analysis, Volume 45.

Motamed, Mohammad and Nobile, Fabio and Tempone, Raúl. (2012). *A Stochastic Collocation Method for the Second Order Wave Equation with a Discontinuous Random Speed.* Numerische Mathematik, Volume 123.