

Short Note

Not peer-reviewed version

A Short Note on Gaussian Distribution with Non-Constant Correlation

[Yudong Tang](#) *

Posted Date: 12 March 2025

doi: 10.20944/preprints202501.0923.v2

Keywords: Gaussian distribution; correlation skew



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Article

A Short Note on Gaussian Distribution with Non-Constant Correlation

Yudong Tang

Independent Researcher, UK; yudong.tang@hotmail.com

Abstract: This article studies the terminal distribution of multi-variate Brownian motion where the correlations are not constant. In particular, with the assumption that the correlation function is driven by one factor, this article developed PDEs to quantify the moments of the conditional distribution of other factors. By using normal distribution and moment matching, we found a good approximation to the true Fokker Planck solution and the method provides a good analytic tractability and fast performance due to the low dimensions of PDEs to solve. This method can be applied to model correlation skew effect in quantitative finance, or other cases where a non-constant correlation is desired in modelling multi-variate distribution.

Keywords: Gaussian distribution; correlation skew

1. Background

Gaussian copula is widely used in quantitative finance modelling. The Gaussian distribution is closely related to an underlying Brownian motion: the standard multi-variate normal distribution is the terminal distribution of an underlying multi-variate Brownian motion where the correlations are constant over time. However the correlation being constant is a limitation of this model which might not fit the actual market. On the other hand, if the correlations are not constant, the result terminal distribution has no closed-form representation in general. Without the closed-form solution or analytic tractability, it becomes less attractive for practical usage. There are research in alternative directions which bypass this tractability issue, for example in [1,2], the respective authors created different terminal distributions which can admit shape with the desired correlation skew effect. In this paper, we still focus on the terminal distribution result from the Brownian motion itself. We study the PDE for the density function and show that with some assumption on the correlation function, we can derive some PDEs which can describe the moments of the marginal and conditional distributions. In particular, these PDEs are of lower dimensions, therefore the calculations are fast and practical. With these moments, we can generate moment matching approximations with nice analytic tractability to the true terminal distribution. The result analytic distribution can be a useful variation to the standard multi-variate normal distribution and it can be used for purpose like modelling correlation skew effect in quant finance.

2. Methodology

2.1. Model Setup

We study this math problem below. This is a 2-dimensional case however we show later that similar techniques can be applied to higher dimensions.

$$x(0), y(0) = 0, 0 \quad (1)$$

$$dx = dw_1 \quad (2)$$

$$dy = \rho(x, y, t)dw_1 + \sqrt{1 - \rho^2(x, y, t)}dw_2 \quad (3)$$

$$< dw_1, dw_2 > = 0 \quad (4)$$

The Fokker-Planck equation [3–5] describes the joint probability density function $p(x, y, t)$ by:

$$\frac{\partial p}{\partial t} = \frac{1}{2} \left(\frac{\partial^2 p}{\partial x^2} + 2 \frac{\partial(\rho p)}{\partial x \partial y} + \frac{\partial^2 p}{\partial y^2} \right) \quad (5)$$

This is a 2d-PDE in the convention of quant finance industry (2d refers to 2-dimension in space variables (x, y) while in fact it is a 3-d PDE if counting t , given the common presence of t in this type of PDE we refer the dimensions to only the space variables) and the general numerical method is slow. However, we can reduce the complexity of the 2d-PDE if we make a reasonable assumption on the correlation function as below:

$$\rho(x, y, t) = \rho(x + y, t) \quad (6)$$

This means the correlation depends on the (x, y) in terms of the total $(x + y)$, which can be interpreted as: correlation depends on a market factor which is the average of the underlyers. With this extra assumption, we can simplify the problem as below:

Lets make change of variables below

$$u = \frac{1}{2}(x + y) \quad (7)$$

$$v = \frac{1}{2}(x - y) \quad (8)$$

Then we have

$$\langle du, dv \rangle = \frac{1}{4} (\langle dx, dx \rangle - \langle dy, dy \rangle) = 0 \quad (9)$$

And du, dv can be written as

$$du = \sqrt{\frac{1 + \rho(u, t)}{2}} dw_3 \quad (10)$$

$$dv = \sqrt{\frac{1 - \rho(u, t)}{2}} dw_4 \quad (11)$$

Note the first diffusion only involves u , then Fokker-Planck equation for probability density of u is a 1d-PDE:

$$\frac{\partial p(u, t)}{\partial t} = \frac{1}{2} \frac{\partial^2}{\partial u^2} \left(\frac{1 + \rho(u, t)}{2} p(u, t) \right) \quad (12)$$

So we can solve $p(u, t)$ first, then we look at the $v(t)$. For any given path $u(s), 0 \leq s \leq t$, the $v(t)$ is simply a sum of infinitesimal normal variables with variances $\frac{1 - \rho(u(s), s)}{2}$, so we know the distribution of $v(t)$ condition on this path $u(s), 0 \leq s \leq t$ is a normal distribution with mean 0 and variance

$$\int_0^t \frac{1 - \rho(u(s), s)}{2} ds \quad (13)$$

2.2. Conditional Distribution and the First Two Moments

Conditioned on a path is not easy to use for calculation, it would be more useful to condition on a value $u(t)$ instead of the whole path. Let $\{v(t)|u(t) = u\}$ be the conditional distribution, note this distribution is not a strict normal distribution in general. A more detailed discussion of this distribution is left to later and now we focus on the first two moments, ie, the mean and variance of this conditional distribution.

The mean is clearly zero by symmetry. For the variance, denoted as $f(u, t)$, we have

Theorem 1. Let $\{v(t)|u(t) = u\}$ be the conditional distribution of $v(t)$ conditioned on $u(t) = u$, then its variance $f(u, t)$ satisfy

$$f(u, t) = E \left[\int_0^t \frac{1 - \rho(u(s), s)}{2} ds | u(t) = u \right] \quad (14)$$

Proof. The proof is straightforward. For completeness included below: By the stochastic integral definition as limit of sums, $v(t)|u(t) = u$ is the following sum with the constraint $u(t) = u$:

$$\sum_i \sqrt{\frac{1 - \rho(u_i, s_i)}{2}} \Delta_i w_4 \quad (15)$$

Using definition of variance, we get this sum

$$\sum_{i,j} \sqrt{\frac{1 - \rho(u_i, s_i)}{2}} \sqrt{\frac{1 - \rho(u_j, s_j)}{2}} \Delta_i w_4 \Delta_j w_4 \quad (16)$$

Which only has non-zero terms as below after taking expectation.

$$\sum_i \frac{1 - \rho(u_i, s)}{2} \Delta_i t \quad (17)$$

with the constraint $u(t) = u$.

Then take the limit of $\Delta_i t \rightarrow 0$ and the statement is proven. $\square$

Note $f(u, t)$ is a path integral on all possible paths $u(s)$ that get to u at t . We have the following:

$$p(u, t + dt)f(u, t + dt) = \int_{-\infty}^{\infty} p(x, t) \left[f(x, t) + dt \frac{1 - \rho(x, t)}{2} \right] p(u, t + dt|x, t) dx \quad (18)$$

The $p(u, t + dt|x, t)$ is the transition probability from state (x, t) to $(u, t + dt)$.

Now we follow the Fokker-Planck equation derivation technique, we will get:

Theorem 2.

$$\frac{\partial}{\partial t}(pf) = p \frac{1 - \rho(u, t)}{2} + \frac{1}{2} \frac{\partial}{\partial u^2} \left(pf \frac{1 + \rho(u, t)}{2} \right) \quad (19)$$

Proof.

$$\begin{aligned} & \frac{\partial}{\partial t}(pf) \\ = & \lim_{dt \rightarrow 0} \frac{p(u, t + dt)f(u, t + dt) - p(u, t)f(u, t)}{dt} \\ = & \lim_{dt \rightarrow 0} \frac{\int_{-\infty}^{\infty} p(x, t)f(x, t)p(u, t + dt|x, t)dx - p(u, t)f(u, t)}{dt} + \lim_{dt \rightarrow 0} \int_{-\infty}^{\infty} p(x, t) \frac{1 - \rho(x, t)}{2} p(u, t + dt|x, t)dx \end{aligned}$$

Note the second term comes to

$$\lim_{dt \rightarrow 0} \int_{-\infty}^{\infty} p(x, t) \frac{1 - \rho(x, t)}{2} p(u, t + dt|x, t)dx = p(u, t) \frac{1 - \rho(u, t)}{2}$$

So we just have to prove

$$\lim_{dt \rightarrow 0} \frac{\int_{-\infty}^{\infty} p(x, t)f(x, t)p(u, t + dt|x, t)dx - p(u, t)f(u, t)}{dt} = \frac{1}{2} \frac{\partial}{\partial u^2} \left(pf \frac{1 + \rho(u, t)}{2} \right)$$

Let $h(u)$ be a smooth function with compact support, consider

$$\begin{aligned} & \int_{-\infty}^{\infty} h(u) \int_{-\infty}^{\infty} p(x, t)f(x, t)p(u, t + dt|x, t)dx du \\ = & \int_{-\infty}^{\infty} p(x, t)f(x, t) \int_{-\infty}^{\infty} h(u)p(u, t + dt|x, t)dudx \\ = & \int_{-\infty}^{\infty} p(x, t)f(x, t) \int_{-\infty}^{\infty} (h(x) + h'(u - x) + \frac{1}{2}h''(u - x)^2 + O((u - x)^3))p(u, t + dt|x, t)dudx \end{aligned}$$

Now the integral $\int_{-\infty}^{\infty} (u-x)^k p(u, t+dt|x, t) du$ is the k -th moment of the Brownian motion $du = \sqrt{\frac{1+\rho(u,t)}{2}} dw_3$, so we have

$$\begin{aligned}\int_{-\infty}^{\infty} (u-x) p(u, t+dt|x, t) du &= 0 \\ \int_{-\infty}^{\infty} (u-x)^2 p(u, t+dt|x, t) du &= \frac{1+\rho(x,t)}{2} dt \\ \int_{-\infty}^{\infty} (u-x)^k p(u, t+dt|x, t) du &= \text{higher order than } dt \text{ when } k > 2\end{aligned}$$

Then we have below, in the order of dt

$$\begin{aligned}& \int_{-\infty}^{\infty} h(u) \int_{-\infty}^{\infty} p(x, t) f(x, t) p(u, t+dt|x, t) dx du \\&= \int_{-\infty}^{\infty} p(x, t) f(x, t) h(x) \int_{-\infty}^{\infty} p(u, t+dt|x, t) du dx + dt \int_{-\infty}^{\infty} \frac{1}{2} p(x, t) f(x, t) h'' \frac{1+\rho(x,t)}{2} dx \\&= \int_{-\infty}^{\infty} p(x, t) f(x, t) h(x) dx + dt \int_{-\infty}^{\infty} \frac{1}{2} p(x, t) f(x, t) h'' \frac{1+\rho(x,t)}{2} dx\end{aligned}$$

so

$$\begin{aligned}& \lim_{dt \rightarrow 0} \frac{\int_{-\infty}^{\infty} h(u) \int_{-\infty}^{\infty} p(x, t) f(x, t) p(u, t+dt|x, t) dx du - \int_{-\infty}^{\infty} h(u) p(u, t) f(u, t) du}{dt} \\&= \int_{-\infty}^{\infty} \frac{1}{2} p(x, t) f(x, t) h'' \frac{1+\rho(x,t)}{2} dx \\&= \int_{-\infty}^{\infty} h(x) \frac{1}{2} \frac{\partial^2}{\partial x^2} (p(x, t) f(x, t) \frac{1+\rho(x,t)}{2}) dx\end{aligned}$$

The last step in above is integration by parts. Because the $h(u)$ is arbitrary smooth function so it follows that:

$$\lim_{dt \rightarrow 0} \frac{\int_{-\infty}^{\infty} p(x, t) f(x, t) p(u, t+dt|x, t) dx - p(u, t) f(u, t)}{dt} = \frac{1}{2} \frac{\partial}{\partial u^2} (p f \frac{1+\rho(u,t)}{2})$$

□

To recap, we have these 2 key equations:

$$\frac{\partial p(u, t)}{\partial t} = \frac{1}{2} \frac{\partial^2}{\partial u^2} \left(\frac{1+\rho(u,t)}{2} p(u, t) \right) \quad (20)$$

$$\frac{\partial}{\partial t} (p f) = p \frac{1-\rho(u,t)}{2} + \frac{1}{2} \frac{\partial}{\partial u^2} (p f \frac{1+\rho(u,t)}{2}) \quad (21)$$

We can solve for p first and then solve for f (It is also possible to bundle the PDE solving for p and f together in discretization etc). Knowing $p(u, t)$ and $f(u, t)$, we know the marginal distribution of u and the mean (0) and variance of the conditional distribution $\{v(t)|u(t) = u\}$. Note we mentioned previously the conditional distribution $\{v(t)|u(t) = u\}$ is not strictly normal in general: A sample of it is basically a two step process: first choose a path for u subject to the $u(t)$ terminal condition, this yields a path-wise integral $\int_0^t \frac{1-\rho(u,s)}{2} ds$. Then choose a point from a normal distribution with variance set to $\int_0^t \frac{1-\rho(u,s)}{2} ds$. Or one can also think it as first choose a sample from a standard normal distribution, then choose a path for u subject to the terminal condition, calculate the $\sqrt{\int_0^t \frac{1-\rho(u,s)}{2} ds}$ and scale the normal variable with it.

One might think the 2nd view can keep the normal ness of the whole sampling result, but actually not: some heriustic thinking is that the normal sampling tends to be centralized, and then the scaling also has some centralized tendency, therefore not the same as a constant scaling will do. This is

heriustic of course, but next we develope equations for higher moment, then one can see it won't be a strict normal as the 4th moment vs 2nd moment relation is different to a normal distribution.

2.3. Higher Order Moments

We can derive the equations of the higher order moments following the similar technique. To demonstrate, we look at the 4-th order. Let $g(u, t)$ be the 4-th order moment of the conditional distribution $\{v(t)|u(t) = u\}$.

Theorem 3.

$$\frac{\partial(pg)}{\partial t} = 3(1 - \rho(u, t))pf + \frac{1}{2} \frac{\partial^2}{\partial u^2} (pg \frac{1 + \rho(u, t)}{2}) \quad (22)$$

Proof. 4-th order moment is limit of sum below with constraint $u(t) = u$

$$\sum_{i,j,k,l} \left(\prod_{m \in \{i,j,k,l\}} \sqrt{\frac{1 - \rho(u_m, s_m)}{2}} \Delta_m w \right) \quad (23)$$

After taking expectation and removing zero terms, this becomes below: note ($i = j$ terms are approaching to 0 when making finer grids so we ignored them)

$$6 \sum_{i < j} \frac{1 - \rho(u_i, s_i)}{2} \frac{1 - \rho(u_j, s_j)}{2} \Delta_i t \Delta_j t \quad (24)$$

Taking limit, in integral representation, it is

$$g(u, t) = E[6 \int_0^t \frac{1 - \rho(u(s), s)}{2} ds \int_s^t \frac{1 - \rho(u(x), x)}{2} dx | u(t) = u] \quad (25)$$

Then following the Fokker Planck derivation, we have:

$$p(u, t + dt)g(u, t + dt) = \int_{-\infty}^{\infty} p(x, t)[g(x, t) + 6dt f(x, t) \frac{1 - \rho(x, t)}{2}] p(u, t + dt | x, t) dx \quad (26)$$

The $p(u, t + dt | x, t)$ is the transition probability from state (x, t) to $(u, t + dt)$. Then following the previous derivation we have

$$\frac{\partial(pg)}{\partial t} = 3(1 - \rho(u, t))pf + \frac{1}{2} \frac{\partial^2}{\partial u^2} (pg \frac{1 + \rho(u, t)}{2}) \quad (27)$$

□

Now we can prove the conditional distribution is not normal in general, because normal distribution's 4th order moment is 3 times of the variance square.

Lemma 1. if $g = 3f^2$, then $\frac{\partial f}{\partial u} = 0$

Proof. Plug into the equations and straightforward. □

If $\frac{\partial f}{\partial u} = 0$ then it becomes the standard multi-variate normal distribution.

2.3.1. Normal Approximation to the Conditional

With above in mind, we still prefer to use the normal distribution with the variance matching $f(u, t)$ to approximate the true conditional distribution. This is because first it can match the moments to the 2nd order which is usually good enough for many practical usage. Secondly, the normal distribution has very good analytic tractability.

Note that the normal distribution is not a random choice either: if we denote the true joint density as $p(u, t)q(u, v, t)$ where $q(u, v, t)$ is the conditional probability of v conditioned on u . Then when we use a normal density form for the q , it will satisfy the Fokker Planck equation on most of the terms.

This leads to a question: is there a good analytic form for the q such that it can match moments to higher order (for example 4-th order)? Obviously such form will involve f, g or the moments it need to match. Due to my limited knowledge, this remains interesting but also a mystery to me.

3. Higher Dimensions

In higher dimensions, similar technique can be applied if we assume the correlations have a dependency on one variable (though the variable might be defined as a linear combination of the base variables) and time only. A brief walk through of the idea as below:

Let $x_1, x_2, \dots, x_n$ be the initial Brownian motion variables with correlations $\rho_{ij}(x_1, x_2, \dots, x_n, t)$. For simplicity and avoid any singularity questions, let's assume the ρ_{ij} all just depend on variable $M = \frac{1}{n} \sum_i x_i$ and t . Now we can represent the random process by new set variables $M, x_2, \dots, x_n$, (x_1 is left out as it can be implied by others). We can do Cholesky decomposition of this set of variables and a nice property is that the Cholesky matrix elements are all just function of M and t : this is because all the ρ_{ij} are just function of M and t and the Cholesky decomposition is a deterministic operation on those ρ_{ij} . Now we can apply the same process as before, solve for the probability density of M , and then for variance function of each of the independent Brownian motion variables coming from the Cholesky decomposition. The process will be long and tedious but there is no conceptual difference to the previous case. Note in a special case where all the $\rho_{ij}(M, t)$ are the same function, the change of variables is much cleaner and easier.

So in general, with the assumption that the correlation dependency degenerated to one variable only, the joint terminal distribution of dimension n can be approximated by solving n 1-d PDEs (1-d refer to 1 space variable).

4. Implementaion Example

We show one example of 2-d case: We discretize p and f together and solve for p first for a time step, and then solve for f . We don't use chain rule to break out the partial derivatives of product but instead discretize on the product. With standard finite difference methods, the calculation is fast and stable. We present an example of the distribution below:

Figure 1 shows the contour of a Gaussian distribution with correlation skew. The underlying correlation function is:

$$\rho(u) = \begin{cases} 0.9 & \text{if } u < -2\sqrt{t} \\ 0.9 - \frac{u+2\sqrt{t}}{4\sqrt{t}}0.4 & \text{if } -2\sqrt{t} \leq u < 2\sqrt{t} \\ 0.5 & \text{if } u \geq 2\sqrt{t} \end{cases}$$

The graph axis is in x and y . Note u, v will be the two diagonal directions.

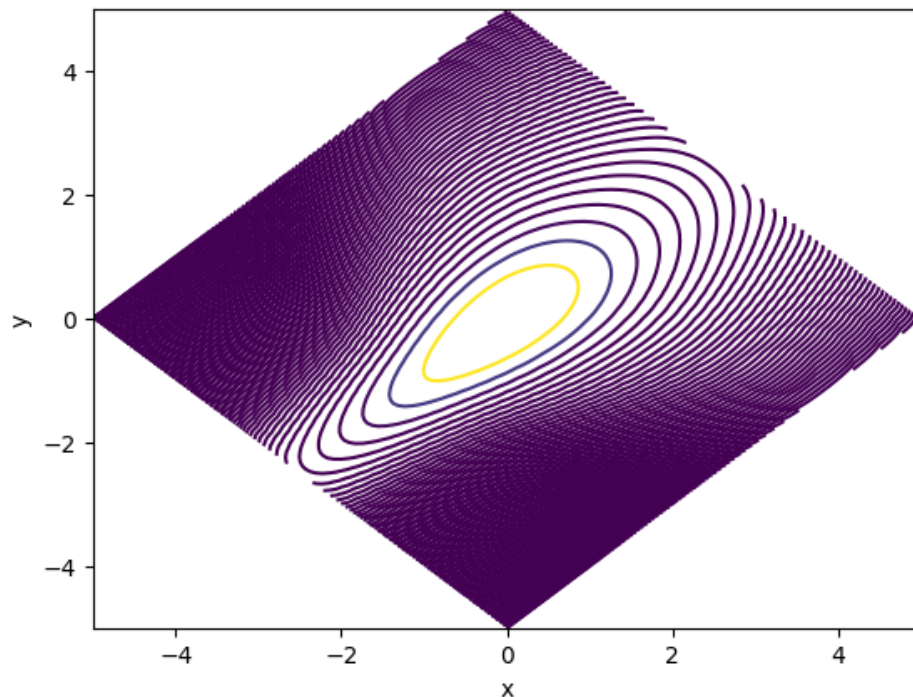


Figure 1. Contour of Gaussian distribution with correlation skew

The shape of the contour is expected. As we put higher correlation when the $u = \frac{x+y}{2}$ is lower, and lower correlation when u is higher, the probability is more concentrated when u is low and more dispersed when u is high. Note with u fixed, the graph also shows symmetry in the direction of v .

The following graphs shows more details on $p(u)$ in above example.

In Figure 2 the distribution of u is very close but different to a standard normal. To see the difference, we reflected the probability around center and then one can see the negative part has a fatter tail than positive part. This is expected as we correlated x, y more when $x + y$ is more negative, we expect $x + y$ will have more potential to go lower in the negative direction, and as we de-correlate x, y more when $x + y$ more positive, we expect the diversifying effect makes the $x + y$ less potential to go higher when $x + y$ positive.

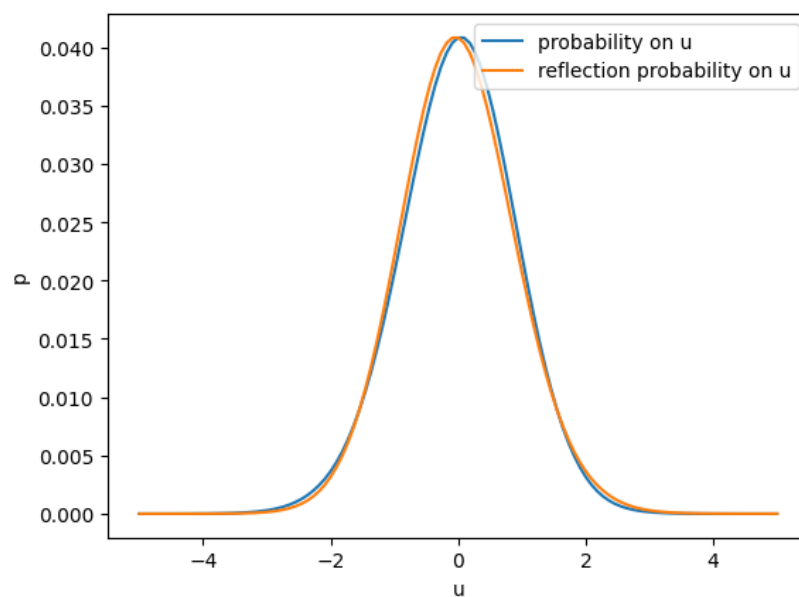


Figure 2. Marginal distribution of u

To demonstrate this point, we can increase the skew of correlation further to see the fat tail effect. Below is the $p(u)$ for a more skewed correlation function.

The correlation function in Figure 3 is

$$\rho(u) = \begin{cases} 0.9 & \text{if } u < -2\sqrt{t} \\ 0.9 - \frac{u+2\sqrt{t}}{4\sqrt{t}} 1.4 & \text{if } -2\sqrt{t} \leq u < 2\sqrt{t} \\ -0.5 & \text{if } u \geq 2\sqrt{t} \end{cases}$$

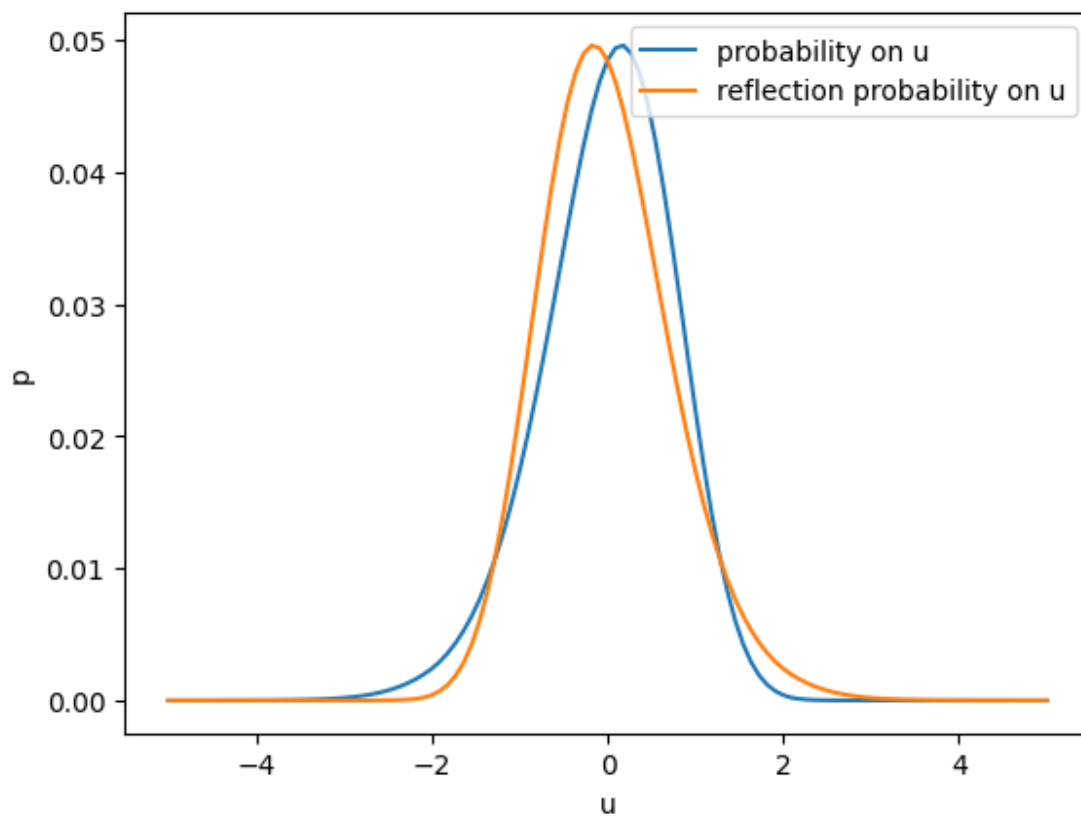


Figure 3. Marginal distribution of u

The contour in Figure 1 shows the v is concentrated when u more negative and v is spreaded when u is more positive. Below Figure 4 shows the std dev of v conditioned on u , ie, the $\sqrt{f}$ function.

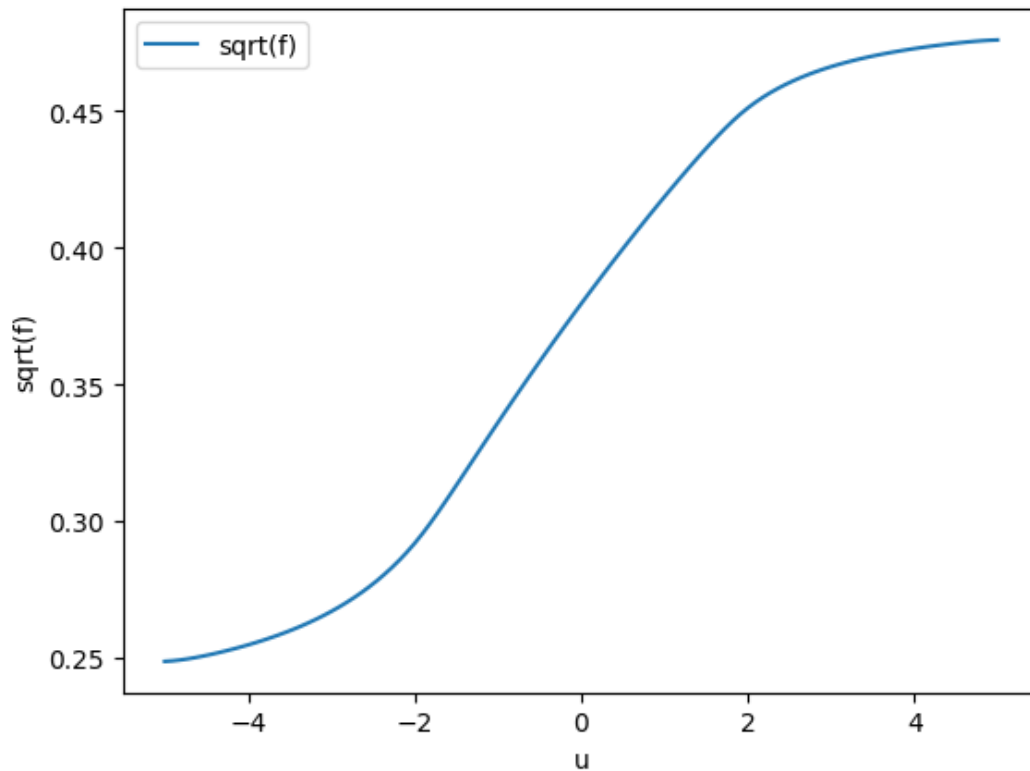


Figure 4. std dev of v conditioned on u

5. Copula Application

Knowing the $p(u)$ and $f(u)$ we can integrate any function on this approximated terminal distribution. For given u, v , it maps to x, y , and the marginal distribution of x and y are approximated normal distribution respectively (Note that the true terminal distribution will have the marginals as true normal distribution, as we approximate the terminal we can think we approximated the marginals), so they can be readily used to invert CDFs.

We found the moment matching to 2nd order produces good accuracy in our tests.

References

1. Lucic, V. (2012). Correlation skew via product copula. In Financial engineering workshop, cass business school.
2. Luj'an, I. (2022). Pricing the correlation skew with normal mean–variance mixture copulas. *Journal of Computational Finance*, 26 (2).
3. Fokker, A. D. (1914). Die mittlere energie rotierender elektrischer dipole im strahlungsfeld. *Annalen der Physik*, 348 (5), 810–820.
4. Planck, V. (1917). Über einen satz der statistischen dynamik und seine erweiterung in der quantentheorie. *Sitzungsberichte der*.
5. Kolmogorov, A. (1931). Über die analytischen methoden in der wahrscheinlichkeitstheorie. *Math Annal*, 104, 415–458.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.