$$f_X(x) = c_k \cdot \sin^k(x),\ x \in (0,\pi),\ k = 1,2,3\ldots \#\text{columns} - 1,\ \text{and}\ c_k = \frac{\Gamma(k/2+1)}{\sqrt{\pi}\,\Gamma(k/2+1/2)}$$

$$F_X(x;k) \sim \frac{1}{2} - \left(\frac{1}{2}\right) \cdot F_{Beta}\left[\cos^2(x); \frac{1}{2}, \frac{1+k}{2}\right] \text{ for } x < \frac{\pi}{2}, \qquad \sim \frac{1}{2} + \left(\frac{1}{2}\right) \cdot F_{Beta}\left[\cos^2(x); \frac{1}{2}, \frac{1+k}{2}\right] \text{ for } x \geq \frac{\pi}{2}$$

The Highly Versatile Angles Space of Positive Definite Dependence Measures: Causal Discovery, Inference, Sampling, and Generalized Entropy

$$F^{-1}(p;k) = \arccos\left(\sqrt{F^{-1}_{Beta}\left(1-2p; \frac{1}{2}, \frac{1+k}{2}\right)}\right) \text{ for } p < 0.5;$$

JD Opdyke
Chief Analytics Officer, DataMineit, LLC
Invited Speaker, PSTAT Seminar,
UCSB Department of Statistics and Applied Probability

$$= \pi - \arccos\left(\sqrt{F^{-1}_{Beta}\left(1-2[1-p]; \frac{1}{2}, \frac{1+k}{2}\right)}\right) \text{ for } p \geq 0.5$$

Acknowledgments & Disclaimer

This presentation contains sections of my forthcoming monograph, “Beyond Correlation: Positive Definite Measures for Robust Inference, Flexible Scenarios, and Causal Modeling for Financial Portfolios.”

Disclaimer: The views presented herein are solely those of the author and do not necessarily reflect the views of specific institutions.

Acknowledgments: I, John Douglas Opdyke, am the sole author of the work contained herein. I would like to sincerely thank Alejandro Rodriguez Dominguez, PhD, for his multiple, invaluable reads of the accompanying monograph, especially for the causal modeling section, that greatly improved both its presentation and its application to capital markets and other empirical research settings. And thanks to Peter Cotton, PhD as well for initially inspiring the application of this work to causal modeling. Any errors are my own.

Abstract

The angles space of positive definite dependence measures possesses many desirable properties that make it useful and of interest for a wide range of empirical inquiry, including: 1. causal discovery; 2. inference for and sampling of these dependence measures; and 3. distance measurement using these measures. Positive definite dependence measures arguably span all those in widespread usage, so this approach maintains a broad and relevant range of application. Angles correspond one-to-one with cells of all-pairwise dependence measure matrices, placing them at the right level of granularity for most analyses of dependence structure. This is as opposed to spectral approaches related to eigen decompositions that remain an order of magnitude less granular, not to mention less robust for estimation when matrices approach singularity (which is the rule rather than the exception for financial portfolios in many settings). Angles distributions are multivariate independent, well bounded, and generally well behaved (i.e. unimodal, not extremely asymmetric, etc.), allowing for reliable estimation and more robust inference. They also remain robust to challenging data conditions, because their only requirement is the positive definiteness of the dependence measure matrix. Real-world financial returns data, for example, is characterized, simultaneously, by marginal distributions with different and varying degrees of asymmetry, heavy-tailedness, serial correlation, and non-stationarity: these conditions in themselves pose no issues for the definition of angle variables or their distributions. Finally, because angles are a multivariate bijection of any positive definite dependence measure matrix, they provide a unifying framework allowing for ceteris paribus comparisons across dependence measures and/or across estimators applied to the same measure, often where no such comparative analyses previously were tractable or even possible. Relying on the angles space, I provide examples of original derivations and empirical implementations of 1., 2., and 3. herein.

Contents

I. Review of Positive Definite Dependence Measures

- I first emphasize that **the dependence structure framework for this work is the all-pairwise matrix**, that is, the matrix of all bivariate relationships between all variables in the portfolio, and that positive definiteness is defined in relation to this framework, NOT in the sense of multivariate measures of dependence (see Cardin, 2009).

$$R = \begin{bmatrix} 1 & r_{1,2} & r_{1,3} & r_{1,4} \\ r_{2,1} & 1 & r_{2,3} & r_{2,4} \\ r_{3,1} & r_{3,2} & 1 & r_{3,4} \\ r_{4,1} & r_{4,2} & r_{4,3} & 1 \end{bmatrix}$$

i. symmetry: $r_{i,j} = r_{j,i}$

ii. unit diagonals: $r_{i=j} = 1$

iii. bounded non-diagonals, maximum range*: $-1 \le r_{i,j} \le 1$

iv. positive definiteness, i.e. all eigenvalues $\lambda_i > 0$

- For completeness, and for reference throughout this presentation, I define eigenvalues for $R \in \mathbb{R}^{p \times p}$ here: if there exists a nonzero vector *v* such that $Rv = \lambda v$ then λ is an eigenvalue of *R* and v is its corresponding eigenvector. λ and *v* can be obtained by solving $\det(\lambda I - R) = 0$, then $\det(\lambda I - R)v = 0$, where I is the identity matrix and det is the determinant.

- Numerical issues aside, **proper measures of dependence in quantitative finance should be positive definite**, either via analytical proof (see Sabado et al., 2007, for Pearson's, Spearman's, and Kendall's, and Embrechts et al., 2016, for the tail dependence matrix), or in the absence of such, via empirical verification of all relevant cases.

*** This range varies by measure, and by data conditions for a given measure, as described in more detail below.**

I. Review of Positive Definite Dependence Measures

Broad Range of Application:

- Even beyond finance, however, positive definite dependence measures arguably span all those in widespread usage, making the use and application of the angles space here broadly relevant.
- For long-established measures (e.g. Pearson's, Spearman's, Kendall's, Tail Dependence Matrix) positive definiteness has been proven analytically.
- But not for newer measures, for which positive definiteness must be empirically verified (e.g. Chatterjee's correlation and its variants (e.g. the improved Chatterjee's correlation (Xia et al., 2024)), Lancaster's correlation(s), Szekely's distance correlation and its variants, Differential Distance Correlation, asymmetric tail dependence measure of Deidda et al. (2023), and others).
- However, because numerical issues can render any estimated dependence measure non-positive definite, such empirical validation always should be employed in practice, regardless of the dependence measure being used.
- Note that all formulae below are estimators based on samples of data, as opposed to populations of data.
- Also note that **this work focuses not on estimation of these dependence measures, but rather on defining and generating their finite sample distributions for inference and related purposes** (e.g. distance measurement, inference for causal discovery, random sample generation). Hence, unless otherwise noted, all results are conditional on 1. a dependence measure matrix, and 2. a data generating mechanism: each can be either specified/known, or well-estimated.

I. Review of Positive Definite Dependence Measures

Monotonic Measures – The "Big 3":

Pearson's Product Moment Correlation

$$1)\ r_{X,Y} = \frac{\sum_{i=1}^{n}\left(X_i - \frac{1}{n}\sum_{j=1}^{n}X_j\right)\left(Y_i - \frac{1}{n}\sum_{j=1}^{n}Y_j\right)\Big/(n-1)}{\sqrt{\sum_{i=1}^{n}\left(X_i - \frac{1}{n}\sum_{j=1}^{n}X_j\right)^2\Big/(n-1)}\sqrt{\sum_{i=1}^{n}\left(Y_i - \frac{1}{n}\sum_{j=1}^{n}Y_j\right)^2\Big/(n-1)}} = \frac{\hat{Cov}(X,Y)}{s_X s_Y}$$

Taking two variables, say, the financial returns of two assets X and Y, Pearson's measures how often and to what degree they deviate from their respective sample means in the same or in opposite directions. The numerator is the (sample) covariance of X and Y, and the denominator – the product of the (sample) standard deviations of X and Y – has the effect of scaling the (sample) covariance to a (maximum) range of –1 to 1,* (from perfect negative dependence to perfect positive dependence, respectively). So Pearson's is just the scaled covariance between X and Y.

* Note that this range can be tighter under specific circumstances, such as for equicorrelation matrices where $\left[-1/(p-1)\right] \le r \le 1,\ p = \dim(r)$.

I. Review of Positive Definite Dependence Measures

Monotonic Measures – The "Big 3":

Spearman's Rho

$$2)\ sr_{X,Y} = \frac{\sum_{i=1}^{n}\left(R_{X_i} - \frac{1}{n}\sum_{j=1}^{n}R_{X_j}\right)\left(R_{Y_i} - \frac{1}{n}\sum_{j=1}^{n}R_{Y_j}\right)\Big/(n-1)}{\sqrt{\sum_{i=1}^{n}\left(R_{X_i} - \frac{1}{n}\sum_{j=1}^{n}R_{X_j}\right)^2\Big/(n-1)}\sqrt{\sum_{i=1}^{n}\left(R_{Y_i} - \frac{1}{n}\sum_{j=1}^{n}R_{Y_j}\right)^2\Big/(n-1)}}$$

Spearman's is exactly the same formula as Pearson's but instead of using the values of the returns of X and Y, their ranks are used instead. It's values also have a maximum range from -1 to 1. In the event of ties in the ranks, widespread convention is to use the averaged rank (see Zar, 1999). If there are no ties in the data, the above can be shortened to 3), although 3) also can be adjusted for ties as in 4), where mx and my = # tied groups in X and Y, respectively, and t_{X_i} and t_{Y_i} are the number of ties in each tied group.

$$3)\ sr_{X,Y} = 1 - \frac{6\sum_{i=1}^{n}\left(R_{X_i} - R_{Y_i}\right)^2}{n^3 - n}$$

$$4)\ sr_{X,Y} = \frac{\left(\left[n^3 - n\right]/6\right) - \sum_{i=1}^{n}\left(R_{X_i} - R_{Y_i}\right)^2 - \sum_{i=1}^{mx}\left(t_{X_i}^3 - t_{X_i}\right)\Big/12 - \sum_{i=1}^{my}\left(t_{Y_i}^3 - t_{Y_i}\right)\Big/12}{\sqrt{\left[\left(\left[n^3 - n\right]/6\right) - 2\sum_{i=1}^{m}\left(t_{X_i}^3 - t_{X_i}\right)\right]\cdot\left[\left(\left[n^3 - n\right]/6\right) - 2\sum_{i=1}^{m}\left(t_{Y_i}^3 - t_{Y_i}\right)\right]}}$$

Computationally, 2) with averaged ranks is preferred to 4) (see Zar, 1999).

I. Review of Positive Definite Dependence Measures

Monotonic Measures – The "Big 3":

Kendall's tau

$$5)\ \tau(X,Y)=\frac{\#\text{concordant pairs}-\#\text{discordant pairs}}{\text{total \# pairs}}=\frac{2}{n(n-1)}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\operatorname{sgn}(x_i-x_j)\operatorname{sgn}(y_i-y_j)$$

$$\text{where } \operatorname{sgn}(z)=1 \text{ if } z>0,\ \operatorname{sgn}(z)=-1 \text{ if } z<0,\ \operatorname{sgn}(z)=0 \text{ if } z=0, \text{ for both } N \text{ and } n$$

Kendall's tau is the sum of all pairwise comparisons of every data point of X and Y, divided by the total number of pairs. The pairwise comparisons are given values of 1, 0, or –1, if both from one period to another are in increasing/decreasing order, if the values from both periods are tied for either of the assets, or if the assets are NOT both in increasing/ decreasing order, respectively; it thus gives the number of pairs in concordance minus the number in discordance relative to the total number of pairs, as shown above in 5).

$$6)\ \tau_b(X,Y)=\frac{1}{\sqrt{(n_0-n_1)(n_0-n_2)}}\left[\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\operatorname{sgn}(x_i-x_j)\operatorname{sgn}(y_i-y_j)\right]$$

Ties in the values of either of the pairs, $(x_i \text{ and } x_j)$ or $(y_i \text{ and } y_j)$, will restrict the range from achieving –1 or +1, even under otherwise perfect discordance or concordance, respectively, so a commonly used variant of Kendall's Tau that avoids this drawback when ties exist is 6).

$$\text{where } n_0=n(n-1)/2;\ n_1=\sum_{i=1}^{i\ \text{grps ties}} t_i(t_i-1)/2;\ n_2=\sum_{j=1}^{j\ \text{grps ties}} u_j(u_j-1)/2;$$

$$t_i=\#\text{ ties in i-th group of } x;\ u_j=\#\text{ ties in j-th group of } y$$

I. Review of Positive Definite Dependence Measures

Monotonic Measures – The "Big 3":

Pearson's rho, Spearman's rho, and Kendall's tau

The "big 3" dependence measures – Pearson's, Kendall's, and Spearman's – are by far the most widely used in practice. Although **widely held myths persist regarding Pearson's as a measure strictly of *linear* monotonic relationships** (see van den Heuvel & Zhan, 2022), all three measure monotonic association (i.e. the direction of the association, positive or negative, does not change within the covered time period) that is symmetric, which is to say non-directional in the variable order (i.e. the measured dependence of X on Y is assumed to be the same as that of Y on X). It is important to recall here that as measures of monotonic dependence, values of zero generally do not imply independence between X and Y,* but independence between X and Y does imply values of zero for the big 3. Some of the dependence measures treated below avoid this limitation under many conditions.

*** Note, however, an exception occurs when data is distributed as bivariate normal, in which case a Pearson's value of zero *does* indicate independence.**

I. Review of Positive Definite Dependence Measures

Tail Dependence Measures:

Tail Dependence Matrix (TDM) provides the probability of a variable exceeding a quantile of its distribution conditional on the other variable in the pair exceeding the same quantile of its distribution. So these conditional probabilities of quantile exceedance range from zero to one, rather than –1 to 1 like the "big 3": otherwise the matrix conditions listed above all hold (its positive definiteness was proven by Embrechts et al., 2016). The upper and lower TDMs only are equal under perfect distributional symmetry: otherwise, the two metrics have distinct values, as shown below in 7) and 8):

$$7)\ TDMU_{X,Y} = \lim_{q \to 1^-} P\left(Y > F_Y^{-1}(q) \mid X > F_X^{-1}(q)\right)$$

$$8)\ TDML_{X,Y} = \lim_{q \to 0^+} P\left(Y \le F_Y^{-1}(q) \mid X \le F_X^{-1}(q)\right)$$

$$\text{where quantile function} = \text{inverse cdf} = F^{-1}(q) = \inf\left\{x \in \mathbb{R} : F(x) \ge q\right\}$$

I. Review of Positive Definite Dependence Measures

Tail Dependence Measures:

Sample estimators for 7) and 8) are presented in Garcin and Nicolas (2023) as 9) and 10):

$$9)\ \hat{\lambda}_{U_{X,Y}}(i/n) = \frac{1-2(i/n)+\hat{C}_n(i/n,i/n)}{1-(i/n)}$$

$$10)\ \hat{\lambda}_{L_{X,Y}}(i/n) = \frac{\hat{C}_n(i/n,i/n)}{(i/n)} \qquad (i/n) = q \text{ for } q < 0.5 \text{ and } (i/n) = (1-q) \text{ for } q \geq 0.5$$

$\hat{F}_{X,n}(x) = \frac{1}{n}\sum_{j=1}^{n}\mathbf{1}\{X_j \leq x\}$ and $\hat{F}_{Y,n}(y) = \frac{1}{n}\sum_{j=1}^{n}\mathbf{1}\{Y_j \leq y\}$ are the empirical cdfs.

$$\hat{C}_n(u,v) = \frac{1}{n}\sum_{j=1}^{n}\mathbf{1}\left\{\hat{F}_{X,n}(X_j) \leq u\right\}\mathbf{1}\left\{\hat{F}_{Y,n}(Y_j) \leq v\right\} \text{ is the empirical copula}$$

These are shown in Schmidt and Stadtmüller (2006) to have good statistical properties (i.e. strong consistency and asymptotic normality).

I. Review of Positive Definite Dependence Measures

Tail Dependence Measures:

Many other measures of tail dependence exist (see AghaKouchak et al., 2013, Babić et al., 2023, Manistre, 2008, Li and Joe, 2024, Krupskii and Joe, 2014, Lauria et al., 2021, and intriguingly, Siburg et al., 2024), but 7) and 8) are the oldest, most widely used, and best understood. Tail dependence is especially important in the risk analytics of financial portfolios because "tail events" often represent the most material financial impacts, are typically associated with non-linear effects and associations, and are closely tied to correlation breakdowns: as is commonly and rightly stated, "when things go bad they go bad together." The phenomenon of "correlation breakdowns" is treated in more detail later in this presentation, but note that the tail dependence matrix has been one of the principal tools used in both the literature and by practitioners to quantitatively estimate it and flag it to mitigate its effects.

I. Review of Positive Definite Dependence Measures

Distance-based and Other New Measures:

Szekely's distance correlation (Szekely et al., 2007) uses two matrices: the matrix of pairwise distances between all X values in the sample, and the same matrix calculated based on all the Y values. To the extent that these matrices vary together, Szekely's will approach 1.0, and to the extent they do not, it will approach zero (which does indicate independence). Its range does not indicate the sign (positive v. negative) of the association. Notably, it can be calculated in arbitrary, and different, dimensions: the sample from X can be drawn, say, from a 3-dimensional distribution, and that from Y can be drawn from 6 dimensions.

First, create n x n distance matrices *a* and *b* by letting

$$a_{i,j} = \left\| x_i - x_j \right\| \text{ and } b_{i,j} = \left\| y_i - y_j \right\|,\ i,j = 1,2,3,...,n \text{ where } \left\| \text{vector } z_n \right\| = \sqrt{z_1^2 + z_2^2 + \cdots + z_n^2}$$

Next, subtract from *a* and *b* their row and column means, and add their respective matrix means, as shown below:

$$A_{i,j} = a_{i,j} - a_{*,j} - a_{i,*} + a_{*,*} \text{ and } B_{i,j} = b_{i,j} - b_{*,j} - b_{i,*} + b_{*,*}$$

Then Szekely's distance correlation is 11):

$$11)\ dcorr = \sqrt{\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} A_{i,j}B_{i,j}} \Bigg/ \sqrt{\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} A_{i,j}^2 \cdot \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} B_{i,j}^2}$$

I. Review of Positive Definite Dependence Measures

Distance-based and Other New Measures:

Lancaster's correlation (see Holzmann and Klar, 2024) – shares several characteristics with Szekely's: its values range from zero to one, a value of zero indicates independence, and it does not indicate with a positive or negative sign whether the dependence between X and Y is positive or negative. Lancaster's correlation in 12) was designed not only to handle non-linear and non-monotonic dependence, but also to improve upon, via increased robustness and generalizability and ease of computation, another dependence measure, the maximal correlation (see Hirschfeld (1935) and Gebelein (1941)).

$$12)\ lan = \max\left(\left|r\left(\tilde{X},\tilde{Y}\right)\right|,\left|r\left(\tilde{X}^2,\tilde{Y}^2\right)\right|\right) \text{ where } \tilde{X} = \Phi^{-1}\left(F_X\left(X\right)\right) \text{ and } \tilde{Y} = \Phi^{-1}\left(F_Y\left(Y\right)\right)$$

where *r* is Pearson's correlation, | | is the absolute value function, Φ^{-1} is the quantile (inverse cdf) function of the standard normal distribution, and *F* is the (empirical) cdf of each variable.

A second version is called linear Lancaster's correlation:

$$13)\ lanL = \max\left(\left|r\left(X,Y\right)\right|,\left|r\left(\breve{X}^2,\breve{Y}^2\right)\right|\right) \quad \text{where } \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \text{ and } \bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

$$\text{and } \breve{X} = \left(X - \bar{X}\right) \Big/ \sqrt{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 \Big/ (n-1)} \text{ and } \breve{Y} = \left(Y - \bar{Y}\right) \Big/ \sqrt{\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2 \Big/ (n-1)}$$

I. Review of Positive Definite Dependence Measures

Asymmetric, Directional Measures:

Unlike the big 3, **Chatterjee's new correlation coefficient** (Chatterjee, 2021) ranges from zero to one asymptotically (it can exceed these bounds slightly under finite samples), and a value of zero does indicate independence. Notably, this is an asymmetric dependence measure, that is, the order of X and Y matters: $\xi_n(X,Y)$ does not necessarily equal $\xi_n(Y,X)$, by design. In other words, the dependence of Y on X is not assumed to be identical to the dependence of X on Y: **dependence is directional**. Such measures are not new, in recent times going back over a dozen years (see Zheng et al., 2012), but they have direct origins in work done at the end of the 19th century (see Yule, 1897, and Allena and McAleerb, 2018, for a thorough review of Yule, 1897).

Rank order X and Y pairs according to X values as $\left(\left(X_{(1)},Y_{(1)}\right),\left(X_{(2)},Y_{(2)}\right),\cdots,\left(X_{(n)},Y_{(n)}\right)\right)$, then

$$14)\ chcorr = \xi_n(X,Y) := 1 - 3\sum_{i=1}^{n-1}|r_{i+1} - r_i| \Big/ \left(n^2 - 1\right) \text{ where } r_i = \text{rank of } Y_i$$

Under ties of X values, break ties uniformly at random, then

$$15)\ chcorr = \xi_n(X,Y) := 1 - n\sum_{i=1}^{n-1}|r_{i+1} - r_i| \Big/ \left(2\sum_{i=1}^{n} l_i\left(n - l_i\right)\right) \text{ where } l_i = \#j \text{ such that } Y_{(j)} \geq Y_{(i)}$$

I. Review of Positive Definite Dependence Measures

Asymmetric, Directional Measures:

Note that Chatterjee's can be made symmetric by simply taking the maximum of two Chatterjee tests, one in each direction:

$$16)\ chcorr\_sym = \max\left[\xi_n(X,Y), \xi_n(Y,X)\right]$$

Importantly, note that the all-pairwise matrix of (bivariate) Chatterjee values – and those of all directional dependence measures – remains symmetric, that is, the value in cell i,j = $\xi_n(X_i, X_j)$ = the value in cell j,i = $\xi_n(X_i, X_j)$. The value in cell j,i $\neq \xi_n(X_j, X_i)$. In other words, the all-pairwise matrix depends on the specific column order of the original data.

I. Review of Positive Definite Dependence Measures

Asymmetric, Directional Measures:

Chatterjee's breakthrough has spawned many variants, such as the "**improved Chatterjee's correlation**" derived by Xia et al. (2024), the motivation of which is to increase power by using inverse distance weightings of all neighboring data values as opposed to just one.

$$17)\ ichcorr = \xi_n^{IM}(X,Y) = 1 - \frac{\sum_{i \neq j}^{n} |r_i - r_j| / |i - j|}{\frac{n+1}{3} \sum_{i \neq j}^{n} |i - j|}$$

"Improved" Chatterjee's also is directional, with values ranging from zero (independence) to one (perfect dependence), and it also can be made symmetric using the maximum of two tests, one in each direction:

$$18)\ ichcorr\_sym = \max\left[\xi_n^{IM}(X,Y), \xi_n^{IM}(Y,X)\right]$$

I. Review of Positive Definite Dependence Measures

Asymmetric, Directional Measures:

Note that 17) typically would be coded with a nested loop, but this can be avoided by calculating the entire equation diagonally, thus increasing speed by an order of magnitude (where an order of magnitude is the dimension of the matrix, p). This can save considerable amounts of real runtime when matrices are large(r) (e.g. p≥100) AND these matrices must be calculated many times in many simulations (e.g. Nsim ≥10,000).

SAS/IML code (v9.4): Improved Chatterjee

```
*** INPUTS: data4chat is raw input returns data, reverse is a flag to reverse column order;
*** OUTPUT ichat is the improved Chatterjee matrix;

 start corr_ichat(data4chat,reverse);
   dim = dimension(data4chat);
   ncols=dim[1,2];
   ncols_m1=ncols-1;
   nrows=dim[1,1];
   nrows_m1=nrows-1;
   ncells=ncols**2;
   ichat=J(ncols,ncols,.);
   ichat[do(1,ncells,ncols+1)]=1;
   if reverse=1 then data4chat=data4chat[,do(ncols,1,-1)];
   if nrows>ncols then do;
     do i=1 to ncols_m1;
       data4chat2=data4chat[,i:ncols];
       call sort(data4chat2);
       do j=i+1 to ncols;
         k=j-i+1;
         rnks = rank(data4chat2[,k]);
         cumsum=0;
         cumwgts=0;
         do q=1 to nrows_m1;
           cumsum=cumsum+sum(abs(dif(rnks,q)/q));
           cumwgts=cumwgts+(nrows-q)/q;
         end;
         ichat[i,j]=1-cumsum/(cumwgts*(nrows+1)/3);
         ichat[j,i]=ichat[i,j];
       end;
     end;
   end;
   else do;
     print "Data provided to corr_ichat subroutine must be rull rank.";
   end;
   return(ichat);
 finish;
```

I. Review of Positive Definite Dependence Measures

Asymmetric, Directional Measures:

Another Chatterjee variant is the measure of Zhang (2024a) that **combines Chatterjee's and Spearman's** in an effort to obtain the best of both worlds: a dependence measure that has reasonable power under cases of non-monotonic, non-linear, and/or cyclical dependence (where Spearman's has little to no power, especially compared to Chatterjee's) as well as reasonable power under monotonic dependence (where Chatterjee's has less power than Spearman's).

$$19)\ zcorrsp = I_{n\_sp}(X,Y) = \max\left\{\left|sr_{X,Y}\right|, \sqrt{5/2}\xi_n(X,Y)\right\}$$

This measure also ranges from zero (independence) to 1.0 (perfect dependence), and can be made symmetric using the maximum function (see Zhang, 2024b):

$$20)\ zcorrsp\_sym = \max\left\{\left|sr_{X,Y}\right|, \sqrt{5/2}\xi_n(X,Y), \sqrt{5/2}\xi_n(Y,X)\right\}$$

I. Review of Positive Definite Dependence Measures

Asymmetric, Directional Measures:

Another measure similar to Chatterjee's is the **Differential Distance Correlation** (DDC) of Liu and Shang (2025). DDC's values range from 0 (independence) to 1.0 (perfect dependence), and like Szekely's distance correlation, DDC can be multidimensional. But when X is univariate so that DDC can be used in an all-pairwise matrix, it is defined as 21) below:

$$21)\ DDC_n(X\,|\,Y) = 1 - \frac{1}{(n-1)}\sum_{i=1}^{n-1}\left\|X_{(i)} - X_{(i+1)}\right\| \Bigg/ \left[\binom{n}{2}^{-1}\sum_{i=1}^{n}(2i-n-1)\,X^{(i)}\right]$$

where $\left\{\left(X_{(i)}, Y_{(i)}\right)\right\}_{i=1}^{n}$ are ordered to satisfy $Y_{(i)} \le \cdots \le Y_{(n)}$ and

$X^{(i)}$ are ordered to satisfy $X_{(i)} \le \cdots \le X_{(n)}$

DDC also can be made symmetric using the maximum function (see Liu, 2025):

$$22)\ DDC_{n-sym}(X,Y) = \max\left[DDC_n(X\,|\,Y), DDC_n(Y\,|\,X)\right]$$

I. Review of Positive Definite Dependence Measures

Asymmetric, Directional Measures:

Directional dependence measures also can be applied only to the tails of X and Y, as correlation breakdowns often are associated specifically with (asymmetric) tail dependence: "Extensive evidence has been gathered showcasing the prevalence of … asymmetric tail interdependence within equity and foreign exchange markets, particularly during times of crisis." (Pramanik, 2024). And "We provide new evidence that lower tail dependence coefficients increased compared to upper ones for all pairs in the COVID-19 crash…" (Ito and Yoshiba, 2025). Below is the **asymmetric tail dependence** measure of Deidda et al. (2023), which is essentially Kendall's Tau applied conditionally, only when the percentile, *q*, of X (or Y) is exceeded:

$$23)\ \hat{\tau}_{X,Y}(q) = \left(2(k-2)!/k!\right) \sum_{1 \le i \le j \le n} \operatorname{sgn}\left(X_i - X_j\right) \operatorname{sgn}\left(Y_i - Y_j\right) I\left(X_i, X_j > X_{(n-k)}\right)$$

where $q = 1 - k/n$, and $k \le n$ is the number of exceedences used in the tail, and *I*() is the indicator function (one when true, zero otherwise) ensuring that only the *k* largest observations of X are used.

I. Review of Positive Definite Dependence Measures

RKHS-based Dependence Measures:

Interestingly, note that measures based on Reproducing Kernel Hilbert Spaces (RKHS), such as HSIC (see Gretton et al., 2007), have been shown by Sejdinovic et al. (2013) to be "precisely the formal extensions of the [Szekely's] distance covariance, where the problem of nonintegrability of weight functions is circumvented by using translation-variant kernels." Additionally, all RKHS-based measures are positive definite, by design, as they are based on the inner product, and so they are included herein.

II. Defining the Angles Space

1. **Narrow but Foundational Case: Pearson's, Identity Matrix, Gaussian Data**
2. Fully General Case: Any Positive Definite Dependence Measure, Any Values, Any Data

Using the angles space, for 1. I derive below the parametric, fully analytic result providing confidence intervals for i) the individual cells of Pearson's correlation matrix, and ii) for the entire matrix, that is, simultaneous confidence intervals.

For 2., I define a sampling algorithm that provides the nonparametric estimate of the finite sample multivariate angles distribution.

1. is presented in a fully interactive spreadsheet at http://www.datamineit.com/DMI_publications.htm

The solution is based on the distribution of the angles *Θ* between the (centered) pair of data vectors. These are related to Pearson's via cosine similarity:

$$\cos(\theta) = \frac{\text{inner product}}{\text{product of norms}} = \frac{\langle \mathbf{X}, \mathbf{Y} \rangle}{\|\mathbf{X}\|\|\mathbf{Y}\|} = \frac{\sum_{i=1}^{N}\left(X_i - E(X)\right)\left(Y_i - E(Y)\right)}{\sqrt{\sum_{i=1}^{N}\left(X_i - E(X)\right)^2}\sqrt{\sum_{i=1}^{N}\left(X_i - E(X)\right)^2}} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \rho, \quad \text{with } 0 \le \theta \le \pi$$

GOAL: Derive the distribution of the angles (because these have desirable properties): then we can transform to angles and back again to the correlation matrix to make inferences about the latter (because the relationship between angles and correlations is a multivariate bijection).

We start with the matrix analogue of cosine similarity, as shown below:

II. Defining the Angles Space

RELATIONSHIP between angles and correlations: Multivariate Bijection

$R =$ p x p Pearson's matrix $= BB^t$ where B is the Cholesky factor of R

$$R = \begin{bmatrix} 1 & r_{1,2} & r_{1,3} & \cdots & r_{1,p} \\ r_{2,1} & 1 & r_{2,3} & \cdots & r_{2,p} \\ r_{3,1} & r_{3,2} & 1 & \cdots & r_{3,p} \\ r_{4,1} & r_{4,2} & r_{4,3} & \cdots & r_{4,p} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ r_{p,1} & r_{p,2} & r_{p,3} & \cdots & 1 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \cos(\theta_{2,1}) & \sin(\theta_{2,1}) & 0 & \cdots & 0 \\ \cos(\theta_{3,1}) & \cos(\theta_{3,2})\sin(\theta_{3,1}) & \sin(\theta_{3,2})\sin(\theta_{3,1}) & \cdots & 0 \\ \cos(\theta_{4,1}) & \cos(\theta_{4,2})\sin(\theta_{4,1}) & \cos(\theta_{4,3})\sin(\theta_{4,2})\sin(\theta_{4,1}) & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \cos(\theta_{p,1}) & \cos(\theta_{p,2})\sin(\theta_{p,1}) & \cos(\theta_{p,3})\sin(\theta_{p,2})\sin(\theta_{p,1}) & \cdots & \prod_{k=1}^{n-1}\sin(\theta_{p,k}) \end{bmatrix}$$

for $i > j$ angles $\theta_{i,j} \in (0, \pi)$.

So from $R = BB^t$, for an individual pairwise correlation we have

$$r_{i,j} = \cos(\theta_{i,1})\cos(\theta_{j,1}) + \sum_{k=2}^{i-1}\left[\cos(\theta_{i,k})\cos(\theta_{j,k})\prod_{l=1}^{k-1}\sin(\theta_{i,l})\sin(\theta_{j,l})\right] + \cos(\theta_{j,i})\prod_{l=1}^{i-1}\sin(\theta_{i,l})\sin(\theta_{j,l}) \quad \text{for } 1 \le i < j \le n$$

And recursively, in the other direction, for an individual angle $\theta_{i,j}$ we have:

For $i > 1$: $\theta_{i,1} = \arccos(b_{i,1})$ for j=1; and $\theta_{i,j} = \arccos\left(b_{i,j} \Big/ \prod_{k=1}^{j-1}\sin(\theta_{i,k})\right)$ for $j > 1$

II. Defining the Angles Space

- **We have a multivariate bijection here: a one-to-one, invertible, non-directional relationship between the matrix of angles and the matrix of correlations**. Transforming back and forth:

  1. estimate the correlation matrix
  2. obtain the Cholesky factorization of the correlation matrix
  3. Use inverse trigonometric functions on 2. to obtain corresponding spherical angles

  And in reverse:

  3. Start with a matrix of spherical angles
  2. apply trigonometric functions to obtain the Cholesky factorization
  1. multiply 2. by its transpose to obtain the corresponding correlation matrix

  see Rapisarda et al. (2007), Rebonato & Jaeckel (2000) and Pourahmadi & Wang (2015) (note a typo in the formula in Pourahmadi & Wang (2015) for the first 3 steps)

- Note for later that **this bijection is true generally, for any positive definite matrix, with any values, under any data.**
- Note the **inverse relationship between angles and correlations**: correlations decrease monotonically in their corresponding angles, i.e. correlations increase as angles decrease to zero, and decrease as angles increase to π (see Zhang et al. (2015) and Lu et al. (2019)). The range from 0 to π rather than 0 to 2π is why this is the p-dimensional hyper-hemisphere rather than the hypersphere.

II. Defining the Angles Space

Correlations to Angles

```
* INPUT rand_R is a valid correlation matrix;

cholfact = T(root(rand_R, "NoError"));
rand_corr_angles = J(nrows,nrows,0);
do j=1 to nrows;
  do i=j to nrows;
    if i=j then rand_corr_angles[i,j]=.;
    else do;
      cumprod_sin = 1;
      if j=1 then rand_corr_angles[i,j]=arcos(cholfact[i,j]);
      else do;
        do kk=1 to (j-1);
          cumprod_sin = cumprod_sin*sin(rand_corr_angles[i,kk]);
        end;
        rand_corr_angles[i,j]=arcos(cholfact[i,j]/cumprod_sin);
      end;
    end;
  end;
end;

* OUTPUT rand_corr_angles is the corresponding matrix of
angles;
```

SAS/IML code (v9.4)

Angles to Correlations

```
* INPUT rand_corr_angles is a valid matrix of correlation angles;

Bs=J(nrows, nrows, 0);
do j=1 to nrows;
  do i=j to nrows;
    if j>1 then do;
      if i>j then do;
        sinprod=1;
        do gg=1 to (j-1);
          sinprod = sinprod*sin(rand_corr_angles[i,gg]);
        end;
        Bs[i,j]=cos(rand_corr_angles[i,j])*sinprod;
      end;
      else do;
        sinprod=1;
        do gg=1 to (i-1);
          sinprod = sinprod*sin(rand_corr_angles[i,gg]);
        end;
        Bs[i,j]=sinprod;
      end;
    end;
    else do;
      if i>1 then Bs[i,j]=cos(rand_corr_angles[i,j]);
      else Bs[i,j]=1;
    end;
  end;
end;
rand_R = Bs*T(Bs);

* OUTPUT rand_R is the corresponding correlation matrix;
```

II. Defining the Angles Space

- For completeness, we include below the definition of the Cholesky factor and corresponding formulae:
- A correlation matrix *R* will be real, symmetric positive-definite, so the unique matrix *B* that satisfies

$$R = BB^T$$

where *B* is a lower triangular matrix (with real and positive diagonal entries), and B^T is its transpose, is the Cholesky factorization of *R*. Formulaically, *B*'s entries are as follows:

$$B_{j,j} = (\pm)\sqrt{R_{j,j} - \sum_{k=1}^{j-1} B_{j,k}^2} \qquad B_{i,j} = \frac{1}{B_{j,j}}\left(R_{i,j} - \sum_{k=1}^{j-1} B_{i,k}B_{j,k}\right) \text{ for } i > j$$

- The Cholesky factorization can be thought of as the matrix analog to the square root of a scalar.

III. Inference and Sampling: A Foundational Special Case

DISTRIBUTION of Angles: Pearson's Under the Gaussian Identity Matrix

- **The Cholesky factorization** of a correlation (positive definite) matrix has rows whose squares sum to 1.0, so it is commonly used as a convenient way to ensure that samples remain on the unit hyper-(hemi)sphere. This **automatically enforces positive definiteness**.

- Pourahmadhi & Wang (2015) and others show that the uniform distribution of positive definite matrices on the p-dimensional hemisphere is proportional to the determinant of the Jacobian, which is defined in terms of the Cholesky factorization as shown below (see also Cordoba et al., 2018)

$$\det\left[J\left(U\right)\right] = 2^{p}\prod_{i=1}^{p-1} u_{ii}^{i} \quad \text{where } U \text{ is the Cholesky factorization of correlation matrix } R = UU^{t}$$

- Pourahmadi & Wang (2015) and others (Makalic and Schmidt, 2018) recognized that under the Gaussian identity matrix, sampling polar/spherical angles based on pdf

$$f_X\left(x;k\right) = c_k \cdot \sin^k\left(x\right),\ x \in \left(0,\pi\right),\ k = 1,2,3,\ldots,\left(\#\text{columns} - 1\right), \text{ and } c_k = \frac{\Gamma\left(k/2+1\right)}{\sqrt{\pi}\,\Gamma\left(k/2+1/2\right)}$$

satisfies this constraint. Although not mentioned in Makalic and Schmidt (2018), importantly note that k = #columns – column# (so for the first column of a p=10x10 matrix, k=9; for the second column, k=8, etc.).

- Note that **the angles distributions are a function of the column number of the matrix** (even though the distributions of the correlations themselves are not).

III. Inference and Sampling: A Foundational Special Case

- For sampling from parametric distributions, ideally the CDF can be inverted analytically, and then **inverse transform sampling** can be used. But here, Makalic and Schmidt (2018) state:

  "Generating random numbers from this distribution is not straightforward as the corresponding cumulative density [sic] function, although available in closed form, is defined recursively and requires O(k) operations to evaluate. The nature of the cumulative density [sic] function makes any procedure based on inverse transform sampling computationally inefficient, especially for large k."

- This turns out not to be the case, as Opdyke (2020, 2022, 2023) derived the analytical CDF of this distribution as shown below:

$$f_X(x;k) = c_k \cdot \sin^k(x),\ x \in (0,\pi),\ k = 1,2,3\ldots\#\text{columns} - 1,\ \text{and}\ c_k = \frac{\Gamma(k/2+1)}{\sqrt{\pi}\,\Gamma(k/2+1/2)}$$

$$F_X(x;k) \sim \frac{1}{2} - c_k \cdot \cos(x) \cdot {}_2F_1\left[\frac{1}{2}, \frac{1-k}{2}; \frac{3}{2}; \cos^2(x)\right] \text{ for } x < \frac{\pi}{2},$$

$$\sim \frac{1}{2} + c_k \cdot \cos(x) \cdot {}_2F_1\left[\frac{1}{2}, \frac{1-k}{2}; \frac{3}{2}; \cos^2(x)\right] \text{ for } x \geq \frac{\pi}{2}$$

$$\text{where the Gaussian hypergeometric* function } {}_2F_1[a,b;c;r] = \sum_n^{\infty} \frac{(a)_n (b)_n}{(c)_n} \cdot \frac{r^n}{n!}$$

$$\text{where } (h)_n = h(h+1)(h+2)\cdots(h+n-1),\ n \geq 1,\ (h)_0 = 1,\ \text{and } |r| < 1,\ c \neq 0,-1,-2,\ldots$$

* Note that the (Gaussian) hypergeometric function is not uncommon in this setting, making an appearance in derivations of the distribution of individual correlations (see Muirhead, 1982, and Taraldsen, 2021), moments of the spectral distribution under some conditions (see Adams et al. 2018, and https://reference.wolfram.com/language/ref/MarchenkoPasturDistribution.html), and the definition of positive definite functions (Franca & Menegatto, 2022).

III. Inference and Sampling: A Foundational Special Case

- However, the hypergeometric function admittedly is unwieldy to use. But this result can be simplified further using two established (if not obscure) identities.

For $c = a+1$ and $0 < r < 1$ simultaneously, which holds in this setting, we have ${}_2F_1[a,b;c;r] = B(r;a,1-b)\left(a/r^a\right)$

where $B(r;a,b) = \int_0^r u^{a-1}(1-u)^{b-1}\,du$ = the incomplete beta function

In addition, we have

$F_{Beta}(r;a,b) = B(r;a,b)/B(a,b)$ where $B(a,b) = \dfrac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ = the complete beta function, so

$B(r;a,b) = F_{Beta}(r;a,b)\cdot B(a,b)$

- Taken together we have:

$$F_X(x;k) \sim \frac{1}{2} - c_k \cdot \cos(x) \cdot F_{Beta}\left[\cos^2(x);\frac{1}{2},\frac{1+k}{2}\right] \cdot \frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{1+k}{2}\right)}{\Gamma\left(\frac{2+k}{2}\right)} \cdot \left([1/2]/\sqrt{\cos^2(x)}\right) \text{ for } x < \frac{\pi}{2},$$

$$F_X(x;k) \sim \frac{1}{2} + c_k \cdot \cos(x) \cdot F_{Beta}\left[\cos^2(x);\frac{1}{2},\frac{1+k}{2}\right] \cdot \frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{1+k}{2}\right)}{\Gamma\left(\frac{2+k}{2}\right)} \cdot \left([1/2]/\sqrt{\cos^2(x)}\right) \text{ for } x \geq \frac{\pi}{2}$$

III. Inference and Sampling: A Foundational Special Case

- Recognizing that the complete Beta function is the inverse of the normalization factor of c(k) for these values, their product equals 1 and cancels, as do the two cosine terms, and we have:

$$F_X(x;k) \sim \frac{1}{2} - \left(\frac{1}{2}\right) \cdot F_{Beta}\left[\cos^2(x); \frac{1}{2}, \frac{1+k}{2}\right] \text{ for } x < \frac{\pi}{2},$$

$$\sim \frac{1}{2} + \left(\frac{1}{2}\right) \cdot F_{Beta}\left[\cos^2(x); \frac{1}{2}, \frac{1+k}{2}\right] \text{ for } x \geq \frac{\pi}{2}$$

- This is a signed Beta distribution. The CDF of the well-known Beta distribution is so straightforward that it readily can be used in a spreadsheet (a link to one is provided below). And now, we can even obtain an analytic* quantile function of the angle distribution:

*** Note that I use the term 'analytic' as opposed to 'closed-form' because I am unaware of a closed-form algorithm for the inverse cdf of the beta distribution (see Sharma and Chakrabarty, 2017, and Askitis, 2017). However, for all practical purposes this is essentially a semantic distinction since this quantile function is hard-coded into all major statistical / econometric / mathematical programming languages.**

III. Inference and Sampling: A Foundational Special Case

Let $p = \Pr(x \geq X)$. Then for $x < \frac{\pi}{2}$,

$$p = \frac{1}{2} - \left(\frac{1}{2}\right) \cdot F_{Beta}\left[\cos^2(x); \frac{1}{2}, \frac{1+k}{2}\right]$$

$$-2p = -1 + F_{Beta}\left[\cos^2(x); \frac{1}{2}, \frac{1+k}{2}\right]$$

$$1 - 2p = F_{Beta}\left[\cos^2(x); \frac{1}{2}, \frac{1+k}{2}\right]$$

$$F_{Beta}^{-1}\left(1 - 2p; \frac{1}{2}, \frac{1+k}{2}\right) = \cos^2(x)$$

$$\sqrt{F_{Beta}^{-1}\left(1 - 2p; \frac{1}{2}, \frac{1+k}{2}\right)} = \cos(x)$$

$$\arccos\left(\sqrt{F_{Beta}^{-1}\left(1 - 2p; \frac{1}{2}, \frac{1+k}{2}\right)}\right) = x$$

Note that arcos is arc-cosine, the inverse of the cosine function.

We must reflect the symmetric angle density for p≥0.5, so we have

$$x = \arccos\left(\sqrt{F_{Beta}^{-1}\left(1 - 2p; \frac{1}{2}, \frac{1+k}{2}\right)}\right) \text{ for } p < 0.5,$$

$$= \pi - \arccos\left(\sqrt{F_{Beta}^{-1}\left(1 - 2[1-p]; \frac{1}{2}, \frac{1+k}{2}\right)}\right) \text{ for } p \geq 0.5$$

III. Inference and Sampling: A Foundational Special Case

- So now we have for **the angles distribution, under the Gaussian identity matrix**, for the first time together, the pdf, cdf, and (analytic) quantile function:

$$f_X(x;k) = c_k \cdot \sin^k(x),\ x \in (0,\pi),\ k = 1,2,3\ldots \#\text{columns} - 1, \text{ and } c_k = \frac{\Gamma(k/2+1)}{\sqrt{\pi}\,\Gamma(k/2+1/2)}$$

$$F_X(x;k) \sim \frac{1}{2} - \left(\frac{1}{2}\right) \cdot F_{Beta}\left[\cos^2(x); \frac{1}{2}, \frac{1+k}{2}\right] \text{ for } x < \frac{\pi}{2},$$

$$\sim \frac{1}{2} + \left(\frac{1}{2}\right) \cdot F_{Beta}\left[\cos^2(x); \frac{1}{2}, \frac{1+k}{2}\right] \text{ for } x \geq \frac{\pi}{2}$$

$$F^{-1}(p;k) = \arccos\left(\sqrt{F^{-1}_{Beta}\left(1-2p; \frac{1}{2}, \frac{1+k}{2}\right)}\right) \text{ for } u < 0.5;$$

$$= \pi - \arccos\left(\sqrt{F^{-1}_{Beta}\left(1-2[1-p]; \frac{1}{2}, \frac{1+k}{2}\right)}\right) \text{ for } p \geq 0.5$$

Importantly, although often ignored in the sampling literature, note that properly adjusting for sample size, *n*, and degrees of freedom gives

$$k \leftarrow k + n - \#cols - 2$$

so

$$k = n - col\# \ - 2$$

- Apparently the first (and only other) presentation of this quantile function result comes from an anonymous blog post in March, 2018, although it was obtained via a different derivation, which serves to further validate the result.

* See Xi'an, March, 2018 (https://stats.stackexchange.com/questions/331253/draw-n-dimensional-uniform-sample-from-a-unit-n-1-sphere-defined-by-n-1-dime/331850#331850 and https://xianblog.wordpress.com/2018/03/08/uniform-on-the-sphere-or-not/).
In the interest of proper attribution, a reference on the website to the book "The Bayesian Choice" hints that the Xi'an pseudonym is Christian Robert, a professor of Statistics at Université Paris Dauphine (PSL), Paris, France, since 2000 (https://stats.stackexchange.com/users/7224/xian).

III. Inference and Sampling: A Foundational Special Case

- Note that contrary to the claims of Makalic and Schmidt (2018), $F^{-1}(p;k)$ now can be used for inverse transform sampling: we simply use random uniform variables with support [0,1] in every cell's $F^{-1}(p;k)$ to obtain a matrix of angles which is then converted to a correlation matrix.
- Note that issues regarding sampling efficiency are central for other algorithms of this case (see Lewandowski et al., 2009, Kurowicka, 2014, Cordoba et al., 2018, and Wang et al., 2018 for some examples; note that Makalic and Schmidt, 2018, use an envelope rejection sampling algorithm).
- **However, those issues now disappear for this method** (Opdyke, 2020, 2022, 2024) because its reliance on the Cholesky factor always places it on the unit hyper-hemisphere, where positive definiteness automatically is enforced.
- To use the naïve approach as an extreme example, merely perturbing non-diagonal correlation values uniformly, between –1 and 1, will generate mostly non-positive definite matrices. Bohn & Hornik (2014) and Pourahmadi & Wang (2015) show that the ratio of valid correlation matrices to all matrices generated this way that LOOK like correlation matrices is

$$\Pr\left(rand\,"R" \sim PosDef\right) = X = \frac{\prod_{j=1}^{p-1}\left[\sqrt{\pi}\,\Gamma\left(\frac{j+1}{2}\right)\right]^{j}}{2^{p(p-1)/2}} < \prod_{j=1}^{p-1}\left[\frac{\sqrt{\pi}}{2}\right]^{j} = \left[\frac{\sqrt{\pi}}{2}\right]^{p(p-1)/2} ;\lim_{p\to\infty}\left[X\right] = 0$$

- **For even relatively small matrices of dimension p=25, the odds of successfully randomly generating a single valid positive definite correlation matrix are less than 2 in ten quadrillion.** Hence, even if some sampling algorithms achieve some efficiency gains, the more appropriate approach is to only sample from a positive definite space to begin with.

III. Inference and Sampling: A Foundational Special Case

- In addition to its straightforward and convenience use of inverse transform sampling, this approach (**Opdyke, 2020, 2022, 2024) also appears to be fastest** (a crown previously held by Makalic and Schmidt, 2018). Rubsamen (2023) compares Makalic and Schmidt's (2018) envelope rejection sampling method to the one derived herein: the latter provides a 30% runtime speedup.

- HOWEVER, **sampling in almost all cases is no longer necessary**, because Opdyke (2020, 2022, 2024) provides exact, finite-sample confidence intervals, both for individual cells of the correlation matrix, and the entire matrix (via simultaneous confidence intervals). How?...

- Because **the angles distributions are multivariate independent** (see see Rapisarda et al., 2007; Tsay and Pourahmadi, 2017; Pourahmadi & Wang, 2015; Zhang et al., 2015; and Ghosh et al., 2021), the CDF of the entire angles matrix is simply the product of the CDFs of all the cells; **the bijection between the angles matrix and the correlation matrix then provides the CDF of the correlation matrix, which is identical to that of the angles matrix**:

$$F_{\theta-matrix}\left(\theta_{i>j}\right)=\prod_{i>j}F_{\theta_{i,j}}\left(a_{i,j}\right) \text{ where } a_{i,j} \text{ is the observed angle value}$$

$$F_{R-matrix}\left(R_{i>j}\right)=\Pr\left(R_{i>j}\le r_{i>j}\right)=\Pr\left(g\left(\theta_{i>j}\right)\le r_{i>j}\right)==\Pr\left(\theta_{i>j}\le g^{-1}\left(r_{i>j}\right)\right)=F_{\theta_{i>j}}\left(g^{-1}\left(r_{i>j}\right)\right)=\prod_{i>j}F_{\theta_{i,j}}\left(a_{i,j}\right)$$

where $g(\ )$ and $g^{-1}(\ )$ are defined as above:

$$R=g\left(\theta\right)=BB^{t}$$

$$\theta_{i>j}=g^{-1}\left(R_{i>j}\right)=:\ \text{for } j=1,\ \theta_{i,1}=\ \arccos\left(b_{i,1}\right);\ \text{for } j>1\ \theta_{i,j}=\ \arccos\left(b_{i,j}\Big/\prod_{k=1}^{j-1}\sin\left(\theta_{i,k}\right)\right)$$

III. Inference and Sampling: A Foundational Special Case

- The multivariate bijection does not directly provide the cdf's of the individual correlation matrix cells (the determinant of the Jacobian is required for this), but for this special case, we still are able to obtain not only the confidence intervals for each cell, but also the (simultaneous) confidence intervals for the entire matrix.
- Because individual correlation distributions are invariant to row/column ordering (see Pourahmadi and Wang, 2025; Lewandowski et al., 2009), and the null is the same for every cell, the correlation values in the upper/lower bound matrix must all have the same value.
- This means that all we have to do is obtain the first angle in cell 2,1, and the cosine of this angle will give the correlation value that must populate the rest of each of the two matrices. For this cell, **angle cdf = correlation cdf, for this column/row only**, because of pairwise cosine similarity.
- Consequently, the cdf values used to obtain the angle in cell 2,1 are simply α/2 and (1- α/2) if we are obtaining the confidence intervals of the correlation cells individually, and $1-\left(1-\alpha/2\right)^{(1/q)}$ and $\left(1-\alpha/2\right)^{(1/q)}$, where q = p(1-p)/2 cells in the matrix, if we are obtaining the confidence interval of the entire matrix. The latter result is due to the fact that the angles are multivariate independent, so the cdf is just the product of the cdf's (level's) of all the cells (this is just "simultaneous confidence intervals" under independence).

III. Inference and Sampling: A Foundational Special Case

Select Summarized Results
(dimension=5x5, n=252)

All Individual Cell Confidence Intervals			Equicorrelation Matrices		Simultaneous (Matrix-Level) Confidence Interval			Equicorrelation Matrices	
α	α/2	1-α/2	Lower	Upper	α	α/2	1-α/2	Lower	Upper
0.20	0.100	0.900	-0.0810	0.0810	0.20	0.100	0.900	-0.1454	0.1454
0.10	0.050	0.950	-0.1039	0.1039	0.10	0.050	0.950	-0.1615	0.1615
0.05	0.025	0.975	-0.1236	0.1236	0.05	0.025	0.975	-0.1761	0.1761
0.01	0.005	0.995	-0.1620	0.1620	0.01	0.005	0.995	-0.2060	0.2060

III. Inference and Sampling: A Foundational Special Case

- Of course, there are infinite combinations of individual cell cdf's that yield this matrix-level cdf, but only this one provides the confidence intervals for the individual cells and the entire matrix simultaneously.
- Both confidence intervals – those of the individual cells and those of the entire matrix – are calculated automatically the interactive excel workbook, where the user can specify sample sizes and α, found at http://www.datamineit.com/DMI_publications.htm
- Both sets of confidence intervals are validated via straightforward simulations of independent Gaussian* multivariate distributions.
- **These are the first analytically defined results for the FINITE SAMPLE quantiles / confidence intervals of Pearson's correlation matrix under the null of the Gaussian* identity matrix.**
- Although not necessary, note in the workbook that the single value for each equicorrelation matrix is validated by pushing the math (angles-to-cdfs, and vice versa) through the entire matrix.
- Note also that the confidence intervals remain subject to the usual mathematical constraints of equicorrelation matrices (e.g. $[-1/(p-1)] \leq r \leq 1.0$ where p is the dimension of the matrix). If this is violated for a given α, larger sample sizes are needed to shrink variance to satisfy this constraint.
- * Finally, note that this result is valid beyond the Gaussian, for some additional distributions: for example, the multivariate uniform.
- We now define the angles distribution for the general case, beyond the (Gaussian) identity matrix, and beyond Pearson's.

IV. Inference and Sampling: The General Case

1. Narrow but Foundational Case: Pearson's, Identity Matrix, Gaussian Data
2. **Fully General Case: Any Positive Definite Dependence Measure, Any Values, Any Data**

- As mentioned above, the multivariate bijection between angles and dependence measure values holds not only for Pearson's, but also for ALL positive definite dependence measures.
- And the multivariate angles distribution remains one characterized by multivariate independence.
- This makes the angles space very useful for inference and sampling.
- We just need to define the angles distribution for the general case.
- The extant literature does not provide an analytic solution here.
- Note that spectral distributions* deviate (dramatically) from Marchenko-Pastur when subjected to **heavy-tailed data** (see Burda et al., 2004, Burda et al., 2006, Akemann et al., 2009; Abul-Magd et al., 2009, Bouchaud & Potters, 2015, Martin & Mahoney, 2018; Heiny and Yao, 2022, and Opdyke, 2022) **and/or serial correlation** (see Burda et al., 2004, 2011, Hisakado and Kaneko, 2023, and Opdyke, 2022). If the analytic solution for angles distributions under general conditions is of similar complexity, it appears to be a non-trivial problem.
- However, this need not be a show-stopper, as we can **estimate the angles distributions non-parametrically, using kernels**.

*** Note that spectral distributions are at the wrong level of aggregation here: one eigenvalue/vector for each of p distributions/variables (say, p=100) is not nearly granular enough for analysis of the p(p-1)/2 (= 4,950) cells.**

IV. Inference and Sampling: The General Case

2. Fully General Case: Any Positive Definite Dependence Measure, Any Values, Any Data

5 Steps for Obtaining Angles Distributions

S1. Simulate N samples (N=10,000 typically is sufficient) based on the dependence measure matrix and the data generating mechanism (each can be either specified/known, or well estimated).

S2. Calculate the corresponding N all-pairwise dependence matrices, and their Cholesky factorizations, and transform each of these factorizations into a lower triangle matrix of angles.

S3. Fit a kernel density to each cell of the matrix of angles based on the N values obtained from the N samples in S2.

S4. Generate N samples based on the kernel densities in S3.

S5. Convert each of the N samples from S4. back to a re-parameterized Cholesky factorization, and then multiply it by its transpose to obtain a set of N validly sampled dependence matrices.

Positive definiteness is enforced automatically as the Cholesky factor places us on the **unit** hyper-hemisphere. **All sample generation of the angles matrix hereafter uses just S4., and S5. if samples of the dependence measure itself are needed.**

IV. Inference and Sampling: The General Case

2. Fully General Case: Any Positive Definite Dependence Measure, Any Values, Any Data

- Comparing the dependence measure matrices generated in Steps S4. and S5. to those generated in Steps S1. and S2. validate the approach, demonstrating that **the angles contain all extant information regarding dependence structure** (see Fernandez-Duren & Gregorio-Dominguez, 2023, and Zhang & Yang, 2023, as well as Opdyke, 2022).
- For mid- to large-sized portfolios/matrices (e.g. p=100+), this comparison also is important for calibrating the bandwidth of the kernels fitted in Step S3., as described below.
- But going forward, we are staying within the angles space, and **conducting inference within the angles space**, and so do not need to use the results from Step S5.

The only difference in the angles distributions now with this general case is that they are defined non-parametrically, rather than parametrically.

Also, the angles distributions are no longer necessarily symmetric, which affects the calculation of two-sided p-values, both for individual angles and for the entire angles matrix, although this remains straightforward as shown below.

IV. Inference and Sampling: The General Case

2. Fully General Case: Any Positive Definite Dependence Measure, Any Values, Any Data

Nonparametric **Kernel Estimation** (S3.) **and Sampling** (S4. and S5.)

- Due to the bounded nature of the angles distributions on $\theta \in (0,\pi)$, the nonparametric kernel must be appropriately reflected at the boundary (see Silverman, 1986) via:

$$\text{if } \theta < 0 \text{ then } \theta \leftarrow -\theta; \text{if } \theta > \pi \text{ then } \theta \leftarrow (2\pi - \theta)$$

- As per the standard implementation, the kernel itself is defined as

$$f_h(\theta) = \frac{1}{N}\sum_{i=1}^{N} K_h(\theta - \theta_i) = \frac{1}{hN}\sum_{i=1}^{N} K_h\left(\left[\theta - \theta_i\right]/h\right)$$

$$\text{Gaussian: } K(\theta) = \left(1/\sqrt{2\pi}\right)\cdot e^{-\theta^2/2} \qquad \text{Epanechnikov: } K(\theta) = (3/4)\cdot\left(1-\theta^2\right),\ |\theta| \le 1$$

- I have tested both the Gaussian and the Epanechnikov kernels extensively in this setting, along with three different bandwidth estimators, *h*, from Silverman (1986) and one from Hansen (2014), respectively: $h = 1.06\cdot\hat{\sigma}\cdot N^{-1/5}$ $h = 0.79\cdot \text{IQR}\cdot N^{-1/5}$ $h = 0.9\cdot\min\left(\text{IQR}/1.34,\hat{\sigma}\right)\cdot N^{-1/5}$ $h = 2.34\cdot\hat{\sigma}\cdot N^{-1/5}$ $\left(\text{Epanechnikov only}\right)$

- As is most settings, the kernel matters much less than the bandwidth, although after extensive usage and testing the best combination here appears to be $h = 0.9\cdot\min\left(\text{IQR}/1.34,\hat{\sigma}\right)\cdot N^{-1/5}$ used with the Epanechnikov kernel. Note that the Epanechnikov kernel is used in very closely related problems in this setting (see Burda and Jarosz, 2022).

IV. Inference and Sampling: The General Case

2. Fully General Case: Any Positive Definite Dependence Measure, Any Values, Any Data

Nonparametric **Kernel Estimation** (S3.) **and Sampling** (S4. and S5.)

- **Bandwidth Calibration**: For larger matrices (e.g. p=100+), bandwidths need to be tightened by multiplying *h* by a factor of 0.15. When there are many cells (e.g. for p=100, #cells=p(p-1)/2=4,950) this tightening avoids a slight cumulative drift in metrics that are aggregated across all the cells (e.g. correlation matrix norms, VaR-based Portfolio Economic Capital, spectral distributions, and LNP (a type of 'generalized entropy' defined in a later section below)).
- Multiplying by this factor for smaller matrices does not adversely affect the density estimation in any way, so this factor always is used.
- For matrices much larger than p=100, a further tightening of this factor may be required, and this is readily determined by empirically comparing the distributions of these aggregated metrics under direct data simulation (Steps S1. and S2.) vs. kernel-based sampling (Steps S4. and S5.).
- Note that this calibration only is a function of the sometimes very large number of cells whose angles distributions need to be estimated: it is unnecessary for small portfolios/matrices.

- Note that algorithms for **sample generation** based on these commonly used kernels (e.g. the Gaussian and Epanechnikov) are widely known. An example of the latter is simply the median of three uniform random variates (see Qin and Wei-Min, 2024).

IV. Inference and Sampling: The General Case

2. Fully General Case: Any Positive Definite Dependence Measure, Any Values, Any Data

Nonparametric **Kernel Estimation** (S3.) **and Sampling** (S4. and S5.)

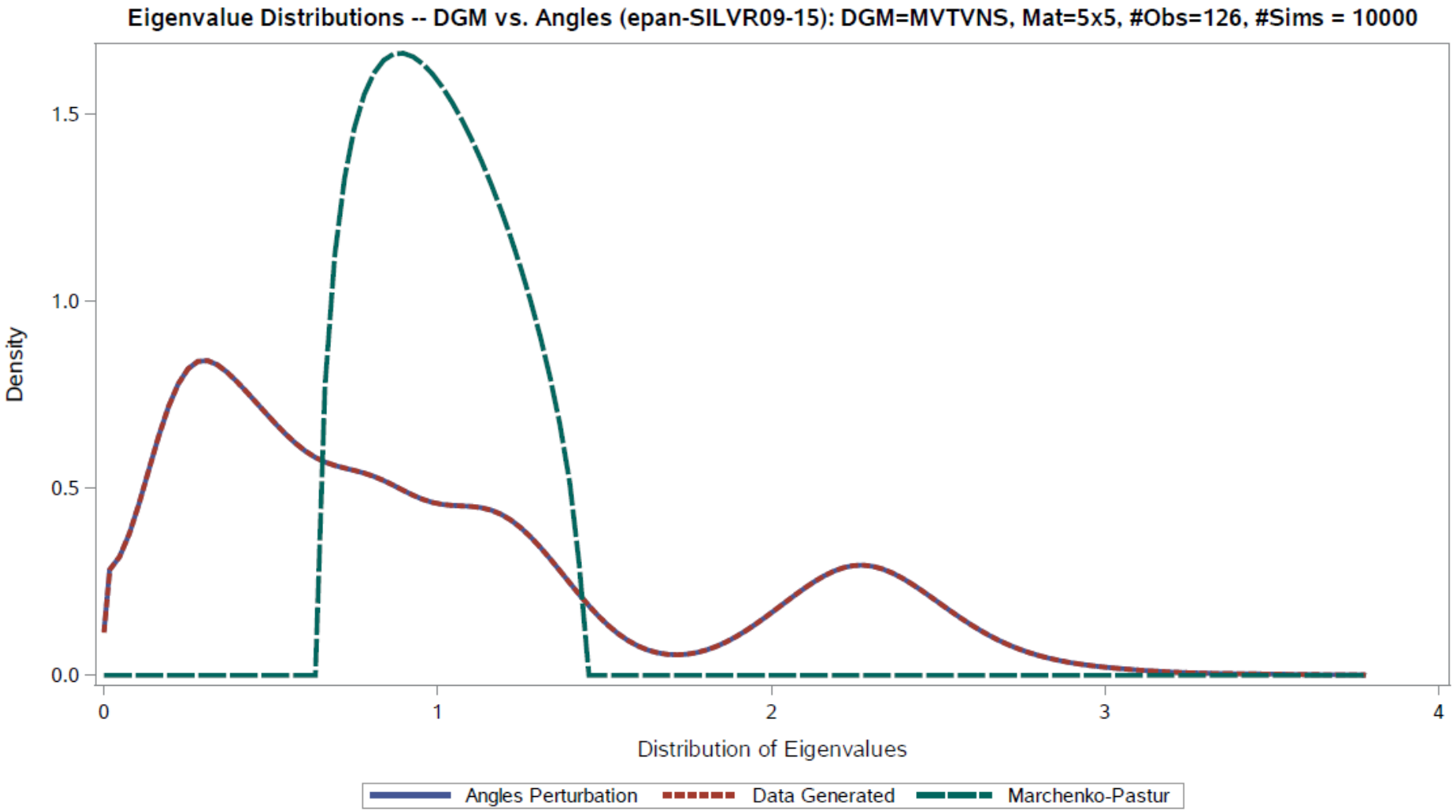
- **Empirical Example**: The angles and spectral graphs below are generated from the **multivariate returns distribution** of the portfolio based on the t-copula of Church (2012),* with p=5 assets, **varying degrees of heavy-tailedness** (df=3, 4, 5, 6, 7), **skewness** (asymmetry parameter=1, 0.6, 0, -0.6, -1), **non-stationarity** (standard deviation=3σ, σ/3, and σ, each with n/3 observations), **and serial correlation** (AR1=-0.25, 0, 0.25, 0.50, 0.75), with a block correlation structure shown below, and n=126 observations for a half year of daily returns.
- The spectral distribution is compared against Marchenko-Pastur as a referential baseline that assumes independence (and identically distributed asset returns with finite variances). Angles distributions also are compared against those from the Gaussian identity matrix.

1	-0.3	-0.3	0.2	0.2
-0.3	1	-0.3	0.2	0.2
-0.3	-0.3	1	0.2	0.2
0.2	0.2	0.2	1	0.7
0.2	0.2	0.2	0.7	1

*** Note that this is only approximately Church's (2012) copula, which incorporates varying degrees of freedom (heavy-tailedness) and asymmetry, because I also impose ex post serial correlation and non-stationarity on the data (and subsequently rescale the marginal densities).**

IV. Inference and Sampling: The General Case

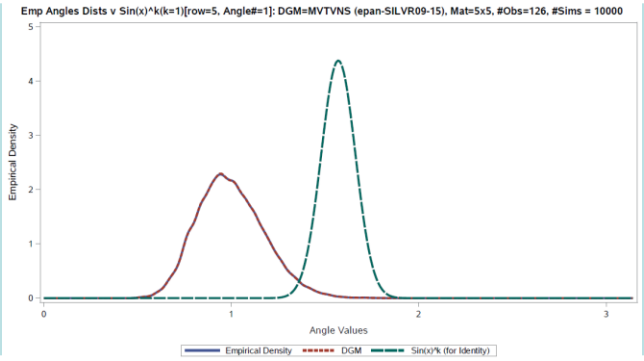
Graph 1a: Spectral Distribution Based on i. Angles Kernel (S4) vs. ii. Direct Data Simulations (S2) vs. Marchenko Pastur

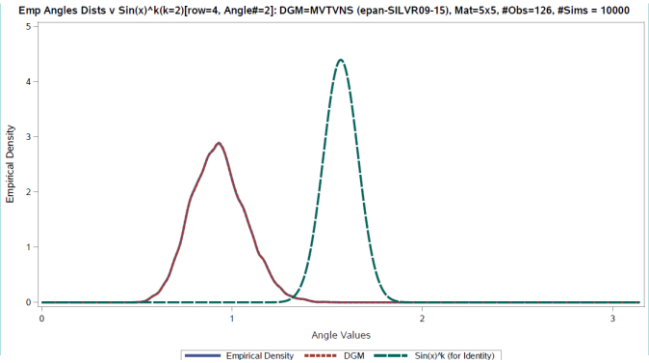


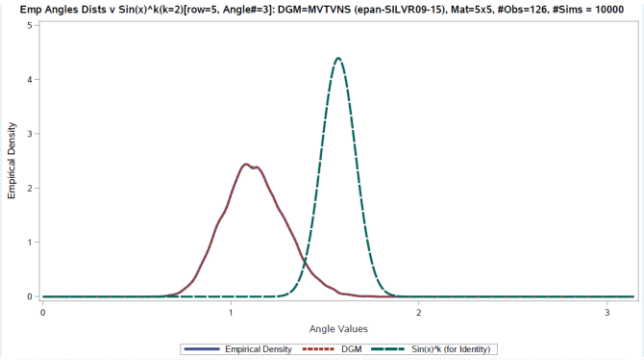
IV. Inference and Sampling: The General Case

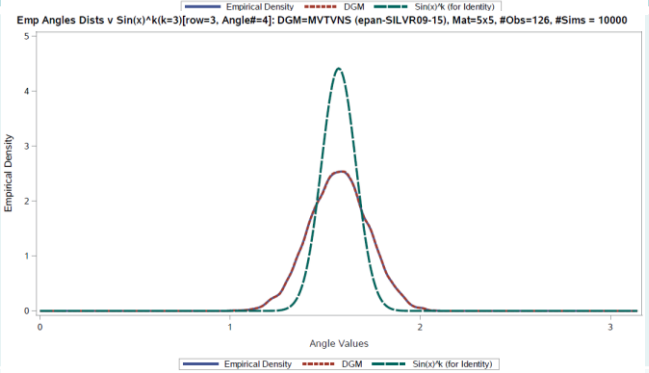
Graphs 1-10:

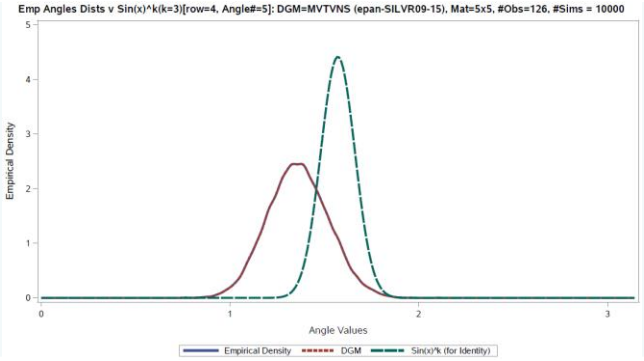
Angles Distributions Based on i. Angles Kernel (S4) vs. ii. Direct Data Simulations (S2) vs. iii. Gaussian Identity Matrix

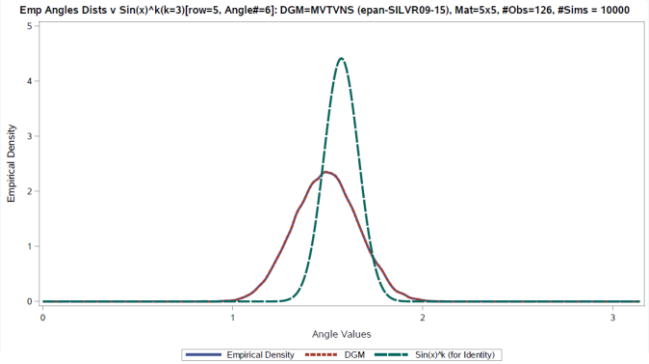








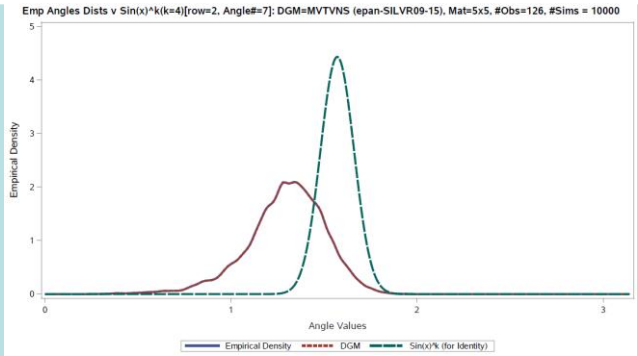


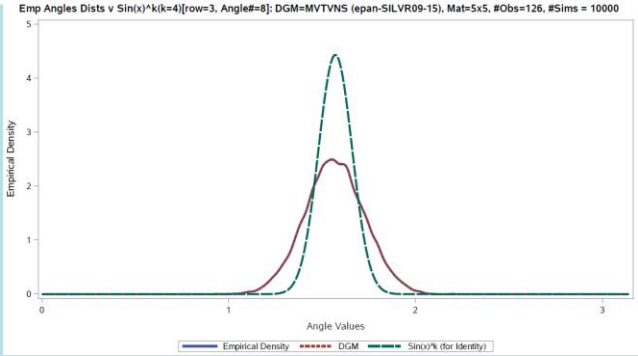


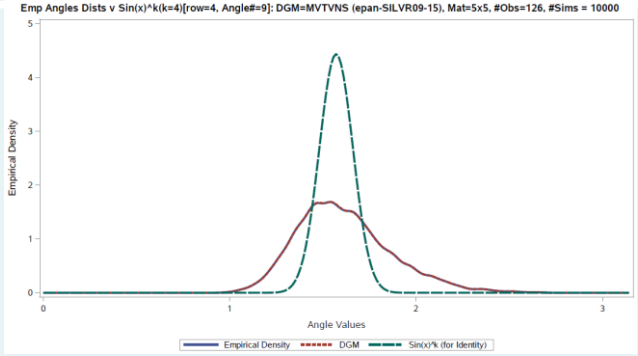
IV. Inference and Sampling: The General Case

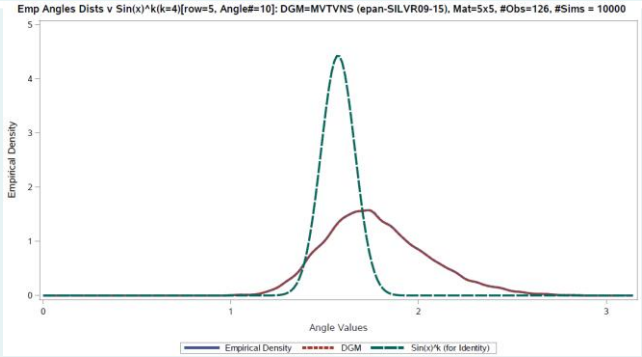
Graphs 1-10:

Angles Distributions Based on i. Angles Kernel (S4) vs. ii. Direct Data Simulations (S2) vs. iii. Gaussian Identity Matrix









IV. Inference and Sampling: The General Case

2. Fully General Case: Any Positive Definite Dependence Measure, Any Values, Any Data

Nonparametric **Kernel Estimation** (S3.) **and Sampling** (S4. and S5.)

- Note that the spectral and angles distributions based on S3.-S5. are essentially identical to those based on S1.-S2., thus **empirically validating the angles sampling approach**.
- Additional aggregated results not presented here (e.g. multivariate VaR-based capital estimation, various norms, and 'generalized entropy' described later below) based on S1.-S2. vs S3.-S5. also are essentially identical, providing further empirical validation.
- Importantly, these results also validate that **the angles contain all extant information regarding dependence structure** (see Fernandez-Duren & Gregorio-Dominguez, 2023, and Zhang & Yang, 2023, as well as Opdyke, 2022).

IV. Inference and Sampling: The General Case

2. Fully General Case: Any Positive Definite Dependence Measure, Any Values, Any Data

Nonparametric **Kernel Estimation** (S3.) **and Sampling** (S4. and S5.)

- Also note that a nonparametric (kernel) approach works in practice here at least in part because the **angles distributions are 'well behaved'**: they are relatively smooth, typically if not always unimodal, and well bounded, so N=10,000 simulations typically suffices to provide a precise and accurate measure of their distributions. This is **in contrast to spectral distributions**, which are spikey, multi-modal, and essentially unbounded for all empirical, practical purposes. These characteristics could require numbers of simulations orders of magnitude larger, which could be computationally prohibitive in larger dimensions (e.g. p=100 with p(p-1)/2=4,950 pairwise associations/cells).
- Also note that on a cell-by-cell basis, the angles distributions deviate materially from each other. There simply is no way that one spectral distribution for a matrix, or even p distributions for each eigenvalue individually, will be able to capture and reflect all the richness of dependence structure reflected here at the granular level of all the p(p-1)/2 pairwise cells. **Eigen decomposition is simply at the wrong level of aggregation to effectively address this more granular, cell-level challenge**, for any real-world purpose in this setting, such as cell-level attribution analyses, granular scenario and reverse scenario analyses, cell-level intervention 'what if' analyses, or customized stress testing, let alone precise and accurate inference.
- Continued reliance on spectral approaches for this problem brings to mind a quotation attributed to John M. Keynes: "the difficulty lies not so much in developing new ideas as in escaping from old ones."

IV. Inference and Sampling: The General Case

2. Fully General Case: Any Positive Definite Dependence Measure, Any Values, Any Data

Nonparametric **Kernel Estimation** (S3.) **and Sampling** (S4. and S5.)

- Going forward below, we conduct inference within the angles space.
- As the angles distributions are no longer necessarily symmetric, care is required when calculating two-sided p-values; however, the calculation remains straightforward (note that all one-sided p-values and confidence intervals automatically take asymmetry into account).
- As before, we are simply using the cdf, and one minus the cdf, to identify both tails.

IV. Inference and Sampling: The General Case

2. Fully General Case: Any Positive Definite Dependence Measure, Any Values, Any Data

Nonparametric **Kernel Estimation** (S3.) **and Sampling** (S4. and S5.)

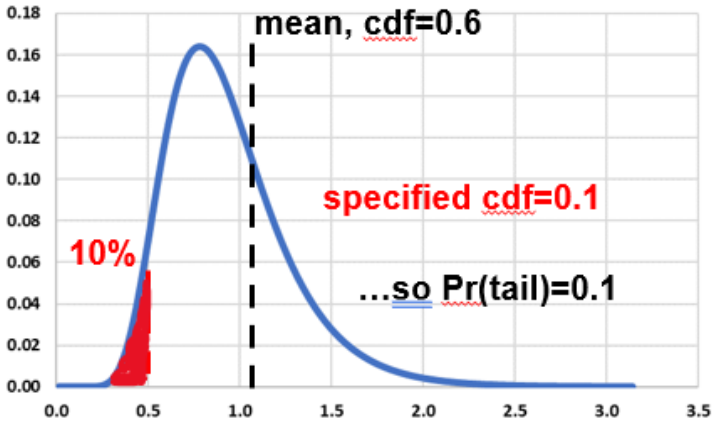
- For a single angle,
  two-sided p-value = max[0, Mcdf – d] + max[0, 1 – (Mcdf + d)], where
  d = abs(Mcdf – cdf), Mcdf = mean angle cdf, cdf = cdf of specified angle
- Here the mean angle cdf (Mcdf) is obtained by i. calculating the mean correlation (dependence measure) matrix across all the simulations; ii. converting it to an angles matrix; iii. determining its cdf's by comparing it to the empirical angles distributions from the simulations.
- So we are only integrating the area in the tails beyond the distance between the specified cell's CDF value and the Mcdf, which under asymmetry is different in each tail.
- If asymmetry is notable, this can sometimes include only one tail, as shown in the graphs below.
- Nevertheless, the p-values in both the one- and two-tailed cases remain 'two-sided' p-values in the statistical sense.

IV. Inference and Sampling: The General Case

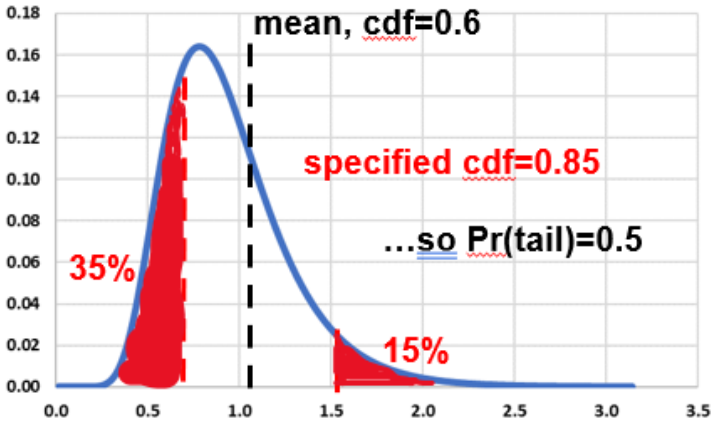
2. Fully General Case: Any Positive Definite Dependence Measure, Any Values, Any Data

Nonparametric **Kernel Estimation** (S3.) **and Sampling** (S4. and S5.)

Graph 11: p-value for a specified angle with a more extreme cdf



Graph 12: p-value for a specified angle with a less extreme cdf



- Note that while cdf=0.1 is hardly more 'extreme' than cdf=0.85 in absolute terms, relative to the mean angle cdf=0.6, it is twice as 'extreme,' i.e. twice as far from the mean cdf=0.6 with a distance of 0.5 for Graph 11, but a distance of only 0.25 for Graph 12.
- Moreover, the tail probability of Graph 11 (0.1) is only 1/5 that of Graph 12 (0.5) (compare the red shaded areas).
- This example demonstrates why asymmetry must be properly taken into account in this setting, but the two-sided p-value still remains a very straightforward calculation, and the "mean angles" matrix is used for additional, important purposes related to scenario restrictions, as shown below.

IV. Inference and Sampling: The General Case

2. Fully General Case: Any Positive Definite Dependence Measure, Any Values, Any Data

Nonparametric **Kernel Estimation** (S3.) **and Sampling** (S4. and S5.)

- Because the angles are multivariate independent, the one-sided p-values of the entire angles matrix are given by:

$$\text{matrix-level (lower) one-sided } p-value = 1-\Pr\left(r_{i>j} \le 0\right) = 1-\Pr\left(a_{i>j} \ge \pi/2\right) = 1-\prod_{i>j}\left(\left[1-F_{i,j}\left(a_{i,j,Mcdf+d}\right)\right]\right)$$

$$\text{matrix-level (upper) one-sided } p-value = 1-\Pr\left(r_{i>j} \ge 0\right) = 1-\Pr\left(a_{i>j} \le \pi/2\right) = 1-\prod_{i>j}F_{i,j}\left(a_{i,j,Mcdf-d}\right)$$

- Similarly, the two-sided p-value for the entire angles matrix is one minus the product of all these 'body,' non-tail probabilities, one for each angle, just as the cdf of the dependence measure matrix is the product of all the cdf's, one for each angle, because the angles are multivariate independent.

$$\text{matrix-level two-sided } p-value = 1-\Pr\left(r_{i>j}=0\right) = 1-\Pr\left(a_{i>j}=\pi/2\right) =$$

$$= 1-\prod_{i>j}\left[1-\left(F_{\theta_{i,j}}\left(a_{i,j,Mcdf-d}\right)+\left[1-F_{\theta_{i,j}}\left(a_{i,j,Mcdf+d}\right)\right]\right)\right]$$

IV. Inference and Sampling: The General Case

2. Fully General Case: Any Positive Definite Dependence Measure, Any Values, Any Data

Nonparametric **Kernel Estimation** (S3.) **and Sampling** (S4. and S5.)

- EMPIRICAL RUNTIMES: Angle-based sampling is relatively computationally intensive, but not prohibitively so. On a commodity laptop purchased in 2024, with RAM=32GB but with no multi-threading, the code processes a 100x100 matrix, with N samples = 10,000 of n = 126 observations each, in about 87 minutes. However, in a multithreaded environment, let alone one with more memory, the angles approach could be applied on similarly non-small matrices much more quickly.
- However, once samples are generated in steps S4. and S5., obtaining correlation matrices for specified cdf's, or obtaining cdf's for specified correlation matrices, or obtaining any related p-values and confidence intervals, is extremely fast, requiring only seconds.
- Applying inverse transform sampling using the analytic quantile function derived herein for the specific case of the Gaussian identity matrix, for a 100x100 matrix, with N samples = 10,000 of n = 126 observations each, takes just over 15 minutes to run on the same laptop.
- When Rubsamen (2023) benchmarked this sampling under the gaussian identity matrix against that of Makalic & Schmidt (2018), he obtained up to a 30% reduction in runtime for the angles sampling derived herein.
- But of course, angles-based inference renders sampling unnecessary under these conditions because it provides the fully analytic solution, as shown above in the fully interactive excel workbook found at http://www.datamineit.com/DMI_publications.htm

IV. Inference and Sampling: The General Case

2. Fully General Case: Any Positive Definite Dependence Measure, Any Values, Any Data

- ESTIMATOR AGNOSTIC: Oftentimes the distributions of specific estimators of dependence measures rely on assumptions or conditions that prohibit ceteris paribus comparisons of the estimators' power, accuracy, robustness, and efficiency. The same is true for the dependence measures themselves, when estimated by the same estimator – the derivations of their distributions rely on different sets of assumptions and conditions, so the methodology acts as a confounder to such all-else-equal comparisons.
- A non-trivially important benefit of the angles-based approach is that **it remains estimator agnostic**: because its only requirement is positive definiteness of the dependence measure, the angles distribution is a common platform wherein such truly all-else-equal comparisons can be performed.
- This is true both when comparing different estimators of the same dependence measure, and when using the same estimator to compare different dependence measures. The power, accuracy, robustness, and efficiency of inferences made based on the angles distribution can be compared, ceteris paribus, under controlled data conditions, with no confounding by methodology-based assumptions or conditions about the distributions of the estimators/measures themselves.
- **The ability to perform either type of ceteris paribus comparison arguably is the exception rather than the rule in the relevant literature on dependence structures: the angles-based approach can in many cases allow the applied researcher to perform both**. Such ceteris paribus comparisons can be performed in the example shown below, where angles-based inference is used for causal discovery.

V. Causal Discovery via DAG Recovery

- The utility of the angles space in analyzing dependence structure indicates a strong potential for its effective application in a very closely related area: **causal model frameworks**.

- **Association-based methods and causal frameworks are not separate modes of inquiry**, somehow at odds with each other, which has made for a convenient but false narrative for some promotional claims (see Domiguez Rodriguez, 2025b, for an analytically rigorous and useful exposé of this): rather, **they often are highly complementary approaches, representing continuous, non-exclusionary improvements in the applied research canon**.

- The facially true, but tired and overly simplistic mantra that “correlation is not causation” unfortunately often is used as a strawman that can do more harm than good for the advancement of causal modeling, because it excludes from consideration some of the most promising approaches for its future development!

- As shown herein, **some of the best causal frameworks make direct and ingenious use of association-based dependence measures**, and this section has been developed with the aim of placing the angles-based approach among these.

V. Causal Discovery via DAG Recovery

- For one example, consider the innovative work of Rodriguez Dominguez (2023, 2025a) in which each portfolio is associated with a common causal manifold, enabling future asset trajectories to be projected into its tangent space. Underlying an important component of this approach, notably, is **the trusty, association-based covariance matrix**.

- Similarly, see the causal model of Cai et al. (2025), as **they utilize the Cholesky factorization of the covariance matrix** as the foundation of their causal algorithm (this use of the Cholesky factor actually is not uncommon; see Li et al., 2025, for another example).

- Additionally, Pascual-Marqui et al. (2024) ingeniously combine Chatterjee's and Szekely's (association-based dependence) measures (shown above) to effectively perform directional, causal regressions.

- Causal modeling is not new, going back at least to Wright (1921), and as we have seen above, neither are **directional measures of association**: in recent times these go back over a dozen years (see Zheng et al., 2012, JASA) but have direct foundations in papers from the nineteenth century (see Yule, 1897, as well as Allena and McAleerb, 2018, for a thorough analysis of Yule, 1897)! **The use of association-based dependence measures within causal model frameworks has made notable inroads** recently (see Pascual-Marqui et al., 2024; Blömbaum et al., 2019; and MacKinnon & Lamp, 2022), **and serves to validate the angles-based approach to DAG recovery described below**.

V. Causal Discovery via DAG Recovery

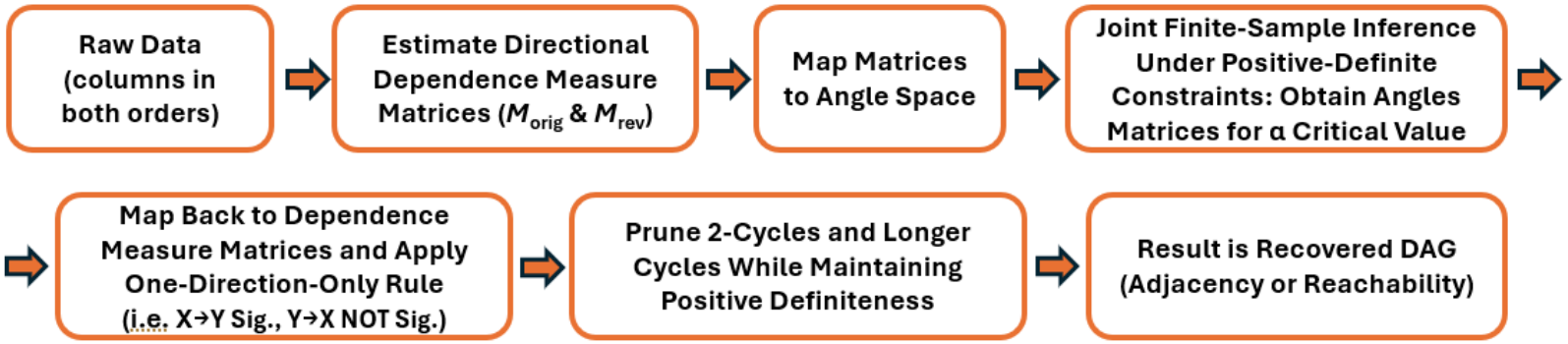
- So given that **association-based models of statistical inference often are foundational to some of the most effective causal frameworks**, I develop below a method using the angles distributions of directional dependence measures as a DAG recovery mechanism.
- Take, for example, Chatterjee's improved correlation ("ICH") coefficient (see Xia et al., 2024), which can be used twice, once with the data variables in one column order, and again with the data variables in the reverse column order. So if we have a treatment variable (X) and dependent variable (Y) and relevant covariates (V1, V2, V3), we column-sort the data in one order (e.g. say, an ordering of X, V1, V2, V3, Y), and then column-sort the data in the reverse order (Y, V3, V2, V1, X), and apply ICH to each ordering; the two resulting ICH matrices will together capture all potential associations, in both directions, of all the variables (factors).
- All the cells of these two estimated dependence matrices will fully map to the relevant variable categories that make up a DAG (e.g. the confounders, colliders, mediators, independent variables, causes of X, consequences of X, causes of Y, and consequences of Y).
- For example, for a particular pair of factors, X and Y (i.e. the X-Y cell of the matrix), two findings from this approach – one of a statistically significant, directional effect of X on Y (X→Y), but NO statistically significant, directional effect of Y on X (Y→X) – provide strong evidence of the directional, causal effect of X on Y (X→Y); the assumptions and definitions provided in the Appendix below subsequently allow for a fully causal interpretation, but I present the rationale for this approach here as follows.

V. Causal Discovery via DAG Recovery

- RATIONALE: Consider a parent–child pair X→Y with $Y = f(X, U_Y)$, $U_Y \perp X$, and NOT Y→X. Under additive-noise mechanisms, independence between X and U_Y implies that the conditional variability of Y given X is dominated by U_Y, whereas the inverse mapping $X = g(Y) + \varepsilon$ generally violates independence (i.e. Y and the induced noise are statistically dependent). Independence between the cause and its exogenous noise term is the identifying restriction that generates directional asymmetry: only the **true causal direction preserves independence, while the inverse mapping generally will violate it. This violation is precisely what directional measures exploit when evaluating the two possible directions of a pairwise relationship.** In other words, because the one direction preserves independence between the input and the associated exogenous noise, while the reverse generally does not, the dual-order calculations allow the expected asymmetry to manifest directly in the empirical directional measures. So we therefore have that directional functionals *M* that reward predictability under independent noise yield $E[M(X \to Y)] > E[M(Y \to X)]$.
- Now **for linear-Gaussian cases, this asymmetry collapses asymptotically as both directions are equivalent. But for nonlinear and/or non-Gaussian and/or heteroskedastic mechanisms, the asymmetry is restored.** And the angles approach leverages this by: (i) computing *M* in both directions (via dual orderings), (ii) performing joint finite-sample statistical tests in angle space under positive definite constraints, and (iii) calling an edge only when one direction is statistically significant and the other direction is not.
- Note that the angles approach is not designed to provide effect sizes, as do regression-based approaches (see for example Pascual-Marqui et al., 2024; Blömbaum et al., 2019; and MacKinnon & Lamp, 2022), nor does it provide counterfactual outcomes. Rather, **the angles approach identifies existing causal structure via DAG recovery**.

V. Causal Discovery via DAG Recovery

GRAPH 19: Angles Approach to DAG Recovery Pipeline

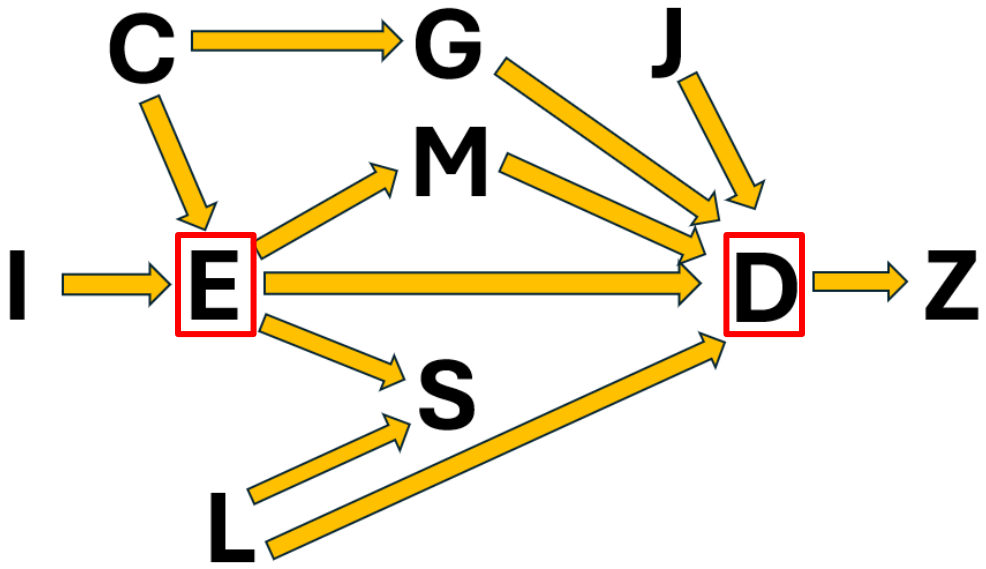


- **The angles approach approach to DAG recovery is original: it calls all edges simultaneously, not sequentially, using directional statistical inference in the angles space wherein positive definiteness automatically is enforced**.
- **This provides potential advantages** over competing algorithms which include:
  a. More coherence in edge calls
  b. Flexibility to use different (directional) dependence measures under different data conditions to increase power
- Both of these potential advantages will be tested against competing algorithms in a more extensive follow-up empirical study.
- For readable exposition here, I place all the necessary definitions and assumptions as they relate to valid causal interpretation in the Appendix (the major assumptions include acyclicity, causal sufficiency, Markov and faithfulness, independent and identically distributed sampling, reliable measurement, and monotone link from (directional) dependence measure to asymmetry).

V. Causal Discovery via DAG Recovery

- For a **preliminary empirical study**, I use the straightforward, but non-toy DAG presented in Digitale et al. (2022) which contains confounders, colliders, and mediators, as well as both causes and consequences of the two primary factors of interest, E and D.

GRAPH 20: Directed Acyclic Graph of Digitale et al., (2022)

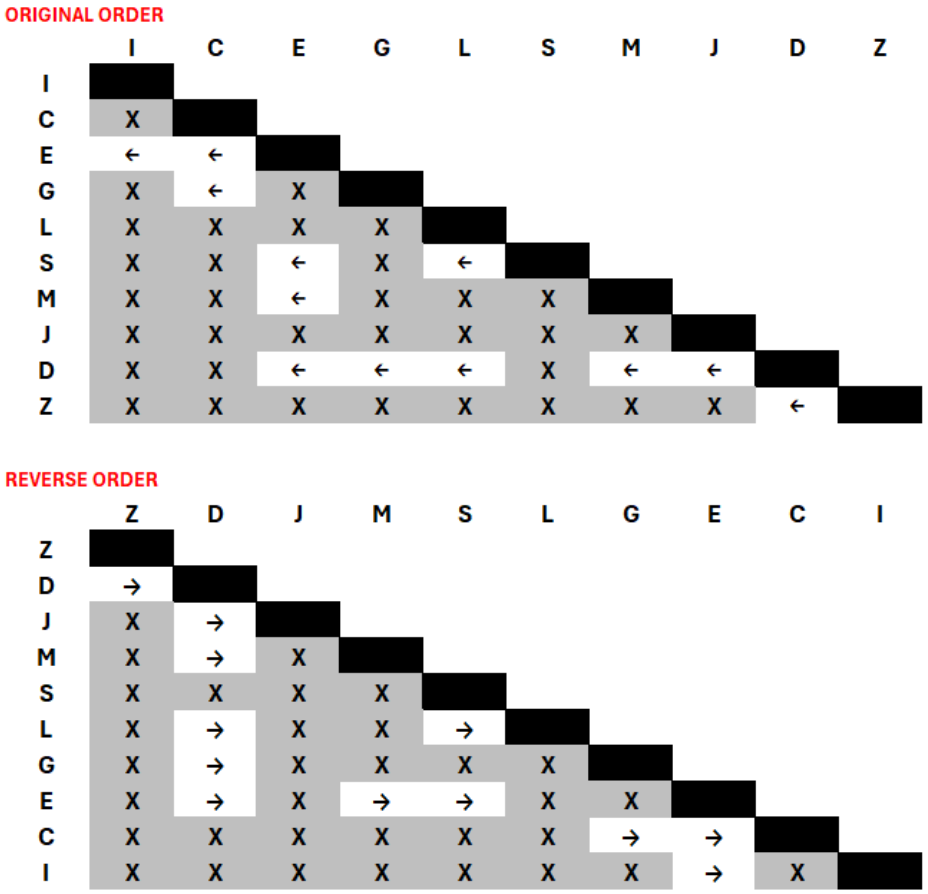


- There are multiple causal paths linking E and D (E→D, E→M→D) as well as multiple non-causal paths linking E and D (E←C→G→D, E→S←L→D). Rightly identifying the directional relationships shown via the arrows in the DAG – and ONLY these relationships – will allow researchers to make unbiased estimation of the effect of E on D (by controlling for other effects on D, but only where appropriate).

V. Causal Discovery via DAG Recovery

- The dimension of the corresponding all-pairwise matrix, where each cell represents an edge of the graph, is p = 10 factors, yielding p(p-1)/2 = 45 pairwise cells. The objective here is to rightly identify the directional effect in each of the 12 cells where one exists, and to rightly identify no effect in the 33 cells that have no effect (see Graph 21).

GRAPH 21: Directional Effects of 12 Pairwise Relationships (Effect Direction is Column to Row, Original Order)



V. Causal Discovery via DAG Recovery

- I generate data under the null hypothesis of independence, to obtain the relevant (one-sided) confidence intervals for the statistical tests on the DAG-generated angles matrices, using multivariate uniform-distributed variables (only independence matters here; distributional changes make virtually no empirical difference in the distributions of the angles). I generate data based on the alternate hypothesis (the DAG) via the code in Table C below.

- **For the DAG data, all parametric distributions are heavy-tailed (student's t-distributed), all error terms are asymmetric (lognormal distributed), and all functional relationships are highly nonlinear (most are sinusoidal, with some quadratic terms)**. As under the null, every simulation has n = 100 observations (which represents about five months of trading days).

- Using reasonably small samples is conservative here, making it more difficult for this approach to rightly identify directional effects, as well as no effects. Strong results under these conditions indicate good power of the estimator and the angles model approach. Note that other studies treating DAG recovery use sample sizes as large as n=1,000 (see for example https://hub.crunchdao.com/competitions/causality-discovery).

- Note that the purpose of the extensive use of coefficients in the DAG code in Table C is only to ensure that the additive effects of multiple parent variables (e.g. those of J, E, M, G, and L on D) are not orders of magnitude different in scale, thus swamping the effects of the other parent variables.

V. Causal Discovery via DAG Recovery

TABLE C: Code Generating Data Under Alternate Hypothesis (DAG)

```
call randgen(R_ERR_LOGN, "LOGNORMAL",0,sqrt(log( (1+sqrt(5))/2 )) );
R_ERR_LOGN = R_ERR_LOGN - exp( log( (1+sqrt(5))/2 )/2 );

call randgen(R_ERR_LOGN2, "LOGNORMAL",0,sqrt(log( (1+sqrt(5))/2 )) );
R_ERR_LOGN2 = R_ERR_LOGN2 - exp( log( (1+sqrt(5))/2 )/2 );

call randgen(R_ERR_LOGN3, "LOGNORMAL",0,sqrt(log( (1+sqrt(5))/2 )) );
R_ERR_LOGN3 = R_ERR_LOGN3 - exp( log( (1+sqrt(5))/2 )/2 );

call randgen(R_ERR_LOGN4, "LOGNORMAL",0,sqrt(log( (1+sqrt(5))/2 )) );
R_ERR_LOGN4 = R_ERR_LOGN4 - exp( log( (1+sqrt(5))/2 )/2 );

call randgen(R_ERR_LOGN5, "LOGNORMAL",0,sqrt(log( (1+sqrt(5))/2 )) );
R_ERR_LOGN5 = R_ERR_LOGN5 - exp( log( (1+sqrt(5))/2 )/2 );

call randgen(R_ERR_LOGN6, "LOGNORMAL",0,sqrt(log( (1+sqrt(5))/2 )) );
R_ERR_LOGN6 = R_ERR_LOGN6 - exp( log( (1+sqrt(5))/2 )/2 );

call randgen(CCC, "T", 10);

call randgen(III, "T",10);

EEE = cos(2*constant('PI')*CCC) + cos(2*constant('PI')*III) + R_ERR_LOGN/8;

GGG = -sin(2*constant('PI')*0.95*CCC) + R_ERR_LOGN2;

call randgen(LLL, "T",10);

SSS = -cos(2*constant('PI')*LLL) - sin(2*constant('PI')*(EEE)) + R_ERR_LOGN3;

MMM = -0.7*cos(2*constant('PI')*(0.95*EEE)) + R_ERR_LOGN4/3;
call randgen(JJJ, "T",10);

DDD = -2.3*cos(2*constant('PI')*(0.6*JJJ)) + 1.35*cos(2*constant('PI')*(0.25*EEE)) - 3.8*MMM##2 +
2.1*cos(2*constant('PI')*(0.65*GGG)) + 2.3*sin(2*constant('PI')*LLL) + R_ERR_LOGN5/8;

ZZZ = cos(2*constant('PI')*(DDD)) + R_ERR_LOGN6/6;
```


V. Causal Discovery via DAG Recovery

- This yields, over 10,000 simulations of the DAG, the average Pearson's, Spearman's, and Kendall's matrices shown in Table D. These show essentially no non-directional associations between the factors/variables.*

TABLE D: Average Pearson's Rho, Spearman's Rho, and Kendall's Tau Matrices
10k Simulations Under the Alternate Hypothesis of the DAG

Pearson's Rho

	I	C	E	G	L	S	M	J	D	Z
I	1									
C	0.001	1								
E	0.001	0.000	1							
G	-0.001	0.001	-0.001	1						
L	-0.001	0.001	0.001	-0.001	1					
S	0.000	0.000	0.050	0.000	-0.001	1				
M	0.001	0.000	0.009	-0.002	0.000	-0.001	1			
J	0.001	-0.001	0.000	0.000	0.000	0.000	0.000	1		
D	-0.001	0.000	0.001	-0.041	0.000	0.001	-0.168	-0.001	1	
Z	0.000	0.001	0.001	-0.001	-0.001	0.001	0.000	-0.001	0.000	1

Spearman's Rho

	I	C	E	G	L	S	M	J	D	Z
I	1									
C	0.001	1								
E	0.001	0.000	1							
G	0.000	0.001	-0.001	1						
L	-0.001	0.001	0.001	-0.001	1					
S	0.000	0.000	0.036	0.000	-0.001	1				
M	0.001	0.000	0.008	-0.001	0.001	-0.001	1			
J	0.001	0.000	0.000	0.000	0.000	0.001	0.000	1		
D	0.000	0.000	0.002	-0.053	-0.001	0.001	-0.039	-0.002	1	
Z	0.000	0.001	0.000	-0.002	-0.001	0.000	0.000	0.000	0.000	1

Kendall's Tau

	I	C	E	G	L	S	M	J	D	Z
I	1									
C	0.000	1								
E	0.000	0.000	1							
G	0.000	0.000	-0.001	1						
L	0.000	0.001	0.000	-0.001	1					
S	0.000	0.000	0.024	0.000	-0.001	1				
M	0.001	0.000	0.005	-0.001	0.000	-0.001	1			
J	0.001	0.000	0.000	0.000	0.000	0.001	0.000	1		
D	0.000	0.000	0.001	-0.036	-0.001	0.001	-0.026	-0.002	1	
Z	0.000	0.001	0.000	-0.001	-0.001	0.000	0.000	0.000	0.000	1

*** One of the cells shows a slight effect, but only under Pearson's (which is more sensitive to outliers), and this is almost certainly due to the different functional form for that factor, which is one (a quadratic term) that is more sensitive to bi-directional effects (when using ICH) than those characterizing most of the DAG (i.e. sinusoidal).**

V. Causal Discovery via DAG Recovery

- I arbitrarily choose a directional dependence measure, the improved Chatterjee correlation coefficient ("ICH") of Xia et al. (2024). As mentioned above, because this is a directional measure whose values range from zero (independence) to one (perfect dependence), I need specifically the upper (1 – α) confidence bound (the upper confidence bound is (1 – α) rather than (1 – α/2) because this is a one-sided hypothesis test) under data generated under independence, to test whether one can reject the null hypothesis of independence when the data are generated under the alternate hypothesis (the DAG).*

- α is set to 0.05, but is conservatively adjusted using a Bonferroni adjustment (α_adj = α/2) because I am conducting two tests on the same data for each edge / cell of the matrix: one with the columns of factor data in the "original" order, and one with the columns of factor data in the reverse order (see Graph 20; the order of the columns in the data match the column order(s) in the matrices).**

*** Recall that inference here is based on the angles corresponding to each cell of the matrix, so this is actually the lower α bound because of the inverse relationship between angles and dependence measure values: angles approach zero as dependence measure values approach one (or perfect dependence), and angles approach π/2 as (directional) dependence measure values approach zero (or independence) (for non-directional measures, like Pearson's or Kendall's or Spearman's, angles approach π as dependence measure values approach negative one, or rather, the lower bound of the matrix, as negative one is a maximal lower bound that is not always attainable).**

**** Regardless of the relationship between these two tests (they are unlikely to be independent), the conservativeness of the Bonferroni adjustment, along with the small number of tests per cell (only two), mitigates any concern about inflated false positives due to multiplicity (see Bonferroni, 1936). The effects of multiplicity across cells (for which the tests are independent, by design, because the angles are independent random variables, as described in earlier sections) are ignored for purposes of this exercise.**

V. Causal Discovery via DAG Recovery

- For each cell, there are four possible outcomes of the two tests:
  A. significant effect in the original direction, no significant effect in the reverse direction
  B. no significant effect in the original direction, significant effect in the reverse direction
  C. no significant effect in either direction
  D. significant effects in both directions

- Under independence, the probabilities associated with A.-D. then become:
  A. Pr(A.) = $\alpha\_adj * (1 - \alpha\_adj) = 0.025*0.975 = 0.024375$
  B. Pr(B.) = $\alpha\_adj * (1 - \alpha\_adj) = 0.025*0.975 = 0.024375$
  C. Pr(C.) = $(1 - \alpha\_adj) * (1 - \alpha\_adj) = 0.975*0.975 = 0.950625$
  D. Pr(D.) = $(\alpha\_adj) * (\alpha\_adj) = 0.025*0.025 = 0.000625$

- The 2-cycles from D. are immediately pruned to preserve the acyclicality of the DAG by retaining only the stronger of the two effects (i.e. the one with the smaller p-value).

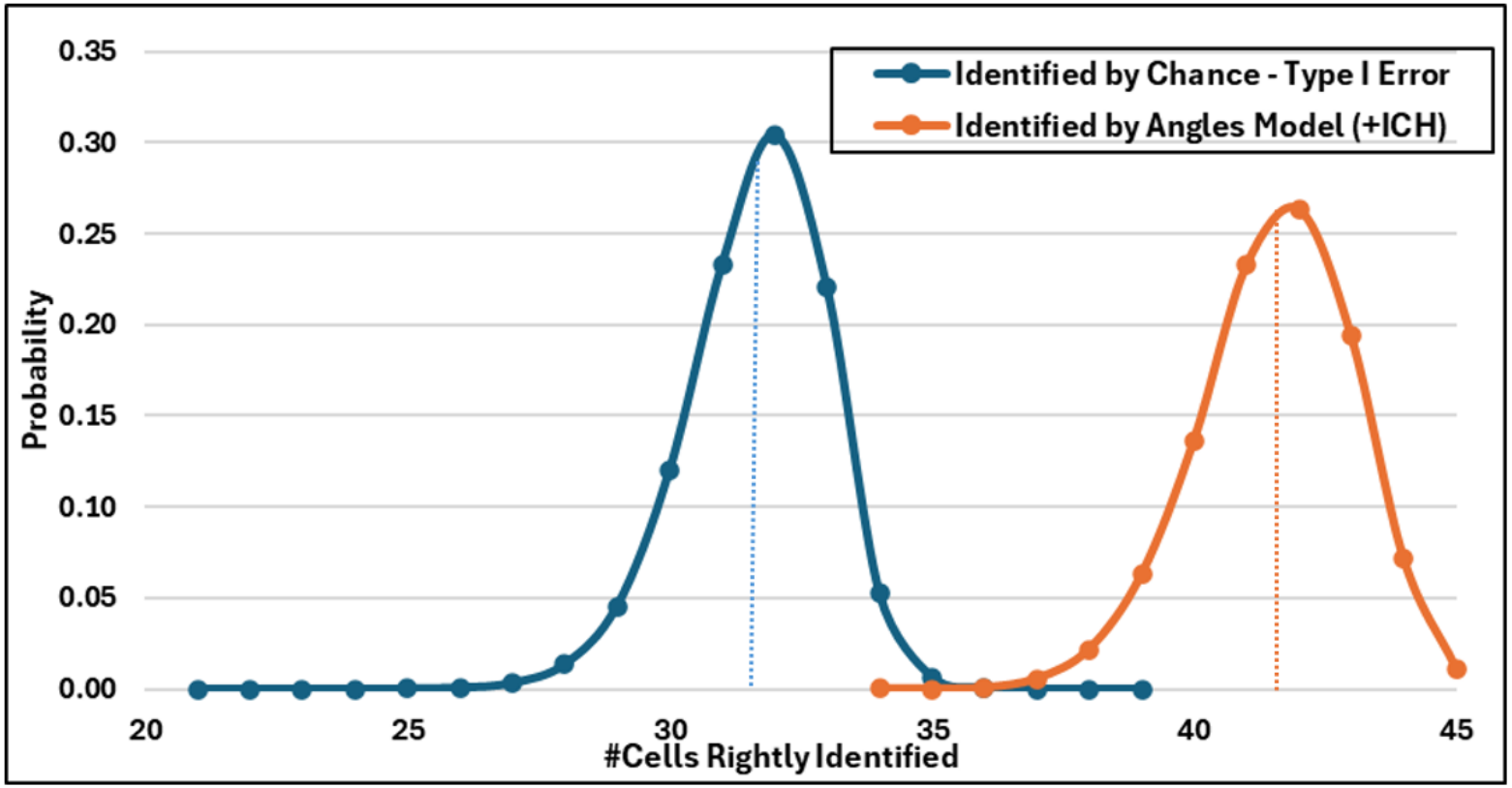
- Next, longer cycles are pruned based on several criteria including the number of called edges (cells) per node (factor/variable), the sum of the p-values associated with all the nodes with the same number of edges, and within each of these groups, the individual p-values of the specific edges (see Appendix for details). All pruning is performed under the constraint of positive definiteness.

- Once **acyclicality is enforced**, what we are interested in is the random variable that is the number of the 45 pairwise cells, which are edges of the DAG, that are rightly identified, where rightly identified means the 12 cells with directional relationships are rightly chosen as A., and those 33 cells with no directional relationships are rightly chosen as C. Under the null hypothesis of independence, this distribution is shown in Graph 22 below, and the expected value is approximately 31.66 right identifications.

V. Causal Discovery via DAG Recovery

- I obtain this null distribution via a simple simulation that uses random uniform variables generated with support [0, 1] to rightly identify each of the 12 directional relationships with Pr(A.) and each of the 33 non-relationships with Pr(C.) yielding, after 100 million simulations, Graph 22, with an expected value of approximately 31.66 right identifications.

GRAPH 22: Distribution of #Rightly Identified Directional Effects/Non-Effects: All 45 Pairs (Fully Exploratory Analysis, No Priors)



V. Causal Discovery via DAG Recovery

- ICH is applied to the data with the factors in original and reverse orders, per Graph 21, and the resulting two matrices of ICH values are saved. The rows and columns of the reverse order matrix are reversed so that the two matrices can be compared easily, cell for cell.

- If, for a given cell, the result of the first test is statistically significant, but the result of the second test is NOT statistically significant, then the result is classified as outcome A.: a directional effect in the right direction. For the 33 'non-relationship' cells, a 'right' outcome is a finding of no statistically significant effect in either direction, i.e. outcome C.

- This is repeated 10,000 times, and for each simulation, acyclicality is enforced, and then the number of times each cell is rightly classified, either as one with a directional effect or one with no effect, is summed and shown in Graph 22 above. The expected value of this distribution is 41.56, compared to that under the null which is 31.66.*

- **The difference between the two distributions, which unarguably is material, also is highly statistically significant**, with a p-value<0.00001.**

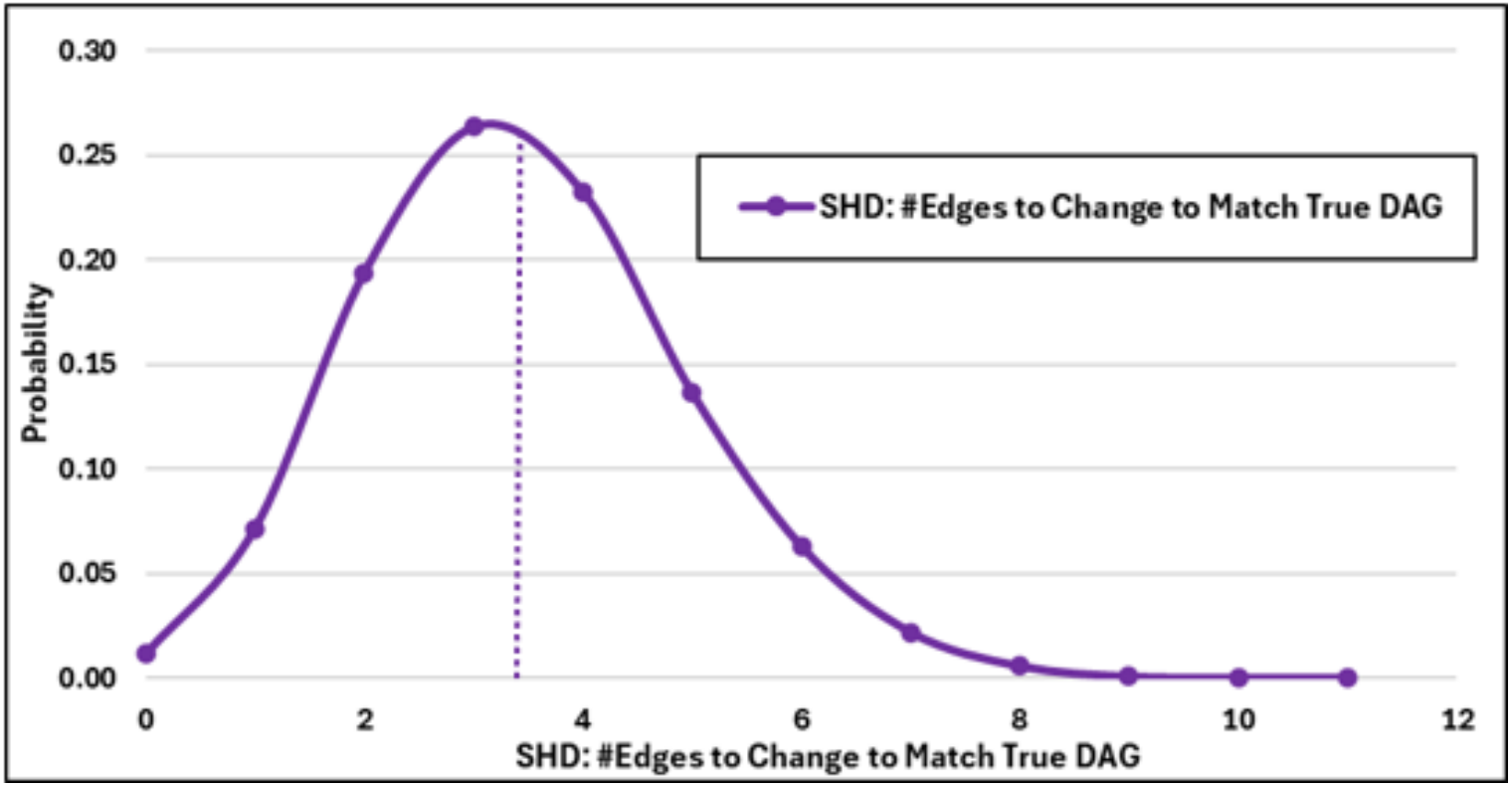
*** Note that acyclicality is not enforced for the distribution under the null, making this a conservative comparison. In other words, the null distribution is likely to be smaller, on average, if edges are pruned when cycles occasionally are formed at random. This is true of the null distributions shown in Graphs 22b, 23, and 24 as well.**

**** This is a one-sided p-value based on a permutation test of 100,000 samples with sample sizes of n1=n2=45 (see Opdyke (2003, 2011, and 2013) for details on permutation tests). These sample sizes are used because they reflect the outcomes of 45 Bernoulli variables, that is, for 45 cells/edges, 45 outcomes of zero or one that sum to the #right in a given sample.**

V. Causal Discovery via DAG Recovery

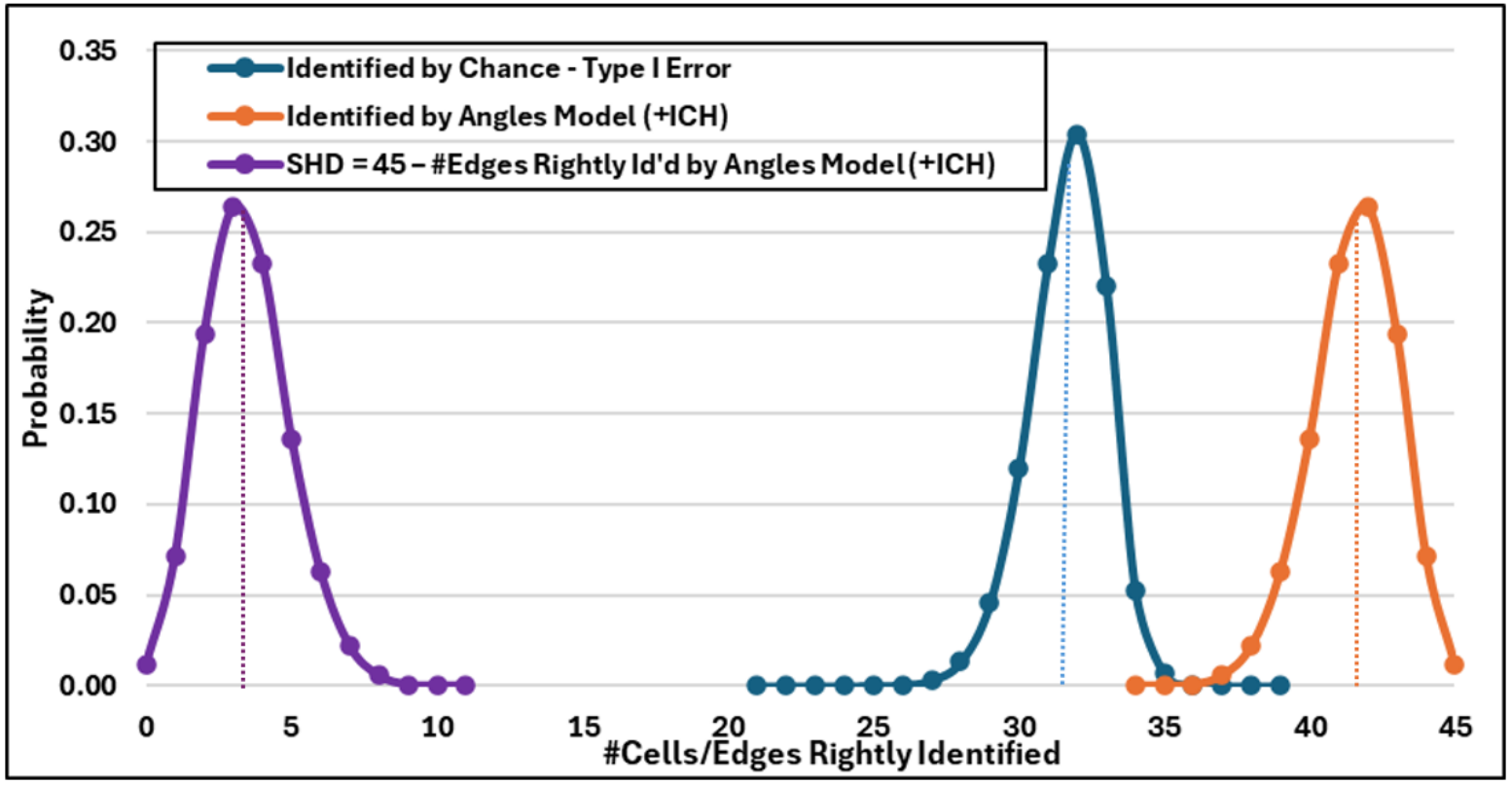
- Note that this **#right metric simply is the inverse of the Structural Hamming Distance (SHD)**, that is, the number of edge modifications that need to be changed for an estimated DAG to match the true DAG. The corresponding SHDs across the 10,000 simulations, with a mean of 3.44, are shown in Graph 22a. I combine Graphs 22 and 22a in Graph 22b below to put them on the same scale.

GRAPH 22a: Distribution of SHD: #Edges Changed to Match True DAG



V. Causal Discovery via DAG Recovery

GRAPH 22b: Distribution of #Rightly Identified Directional Cells/Edges vs. SHD



- In addition to SHD, I present the confusion matrix, otherwise known as the error matrix, below in Table E, along with various common model performance metrics based on it. These values all indicate **strong predictive power of the angles-based model under these conditions**.

V. Causal Discovery via DAG Recovery

TABLE E: Error Matrix, and Associated Performance Metrics

Error Matrix - Numbers

	Predicted: Right	Predicted: Not Right	TOTAL
Right Direction	94,694	25,306	120,000
Not Right Direction	9,127	320,873	330,000

Error Matrix - Rates

	Predicted: Right	Predicted: Not Right	TOTAL
Right Direction	78.9%	21.1%	100%
Not Right Direction	2.8%	97.2%	100%

Precision =	0.91	MCC =	0.80	Jaccard =	0.73
Recall =	0.79	F1 =	0.85	Prevalence Threshold=	0.16
Accuracy =	0.92	Youden =	0.76	Balanced Accuracy =	0.88

V. Causal Discovery via DAG Recovery

- Table F presents the %right classifications, across all simulations, for each specific cell / edge. Red font indicates lower relative scores on the %right metric. For the effect of M on D the red is almost certainly due to the functional form being a quadratic term as opposed to a sinusoidal term, where the former is more sensitive to bi-directional results (when using ICH) and thus, yields fewer rightly identified one-direction effects; also, the effect of E on D is (relatively) muted due to the large number of parent nodes to D. This type of table can be very informative regarding the **specific strengths and weaknesses of an estimator and its framework under different DAGs and different data conditions**.

TABLE F: %Right Relationships Identified in Each Cell / Edge

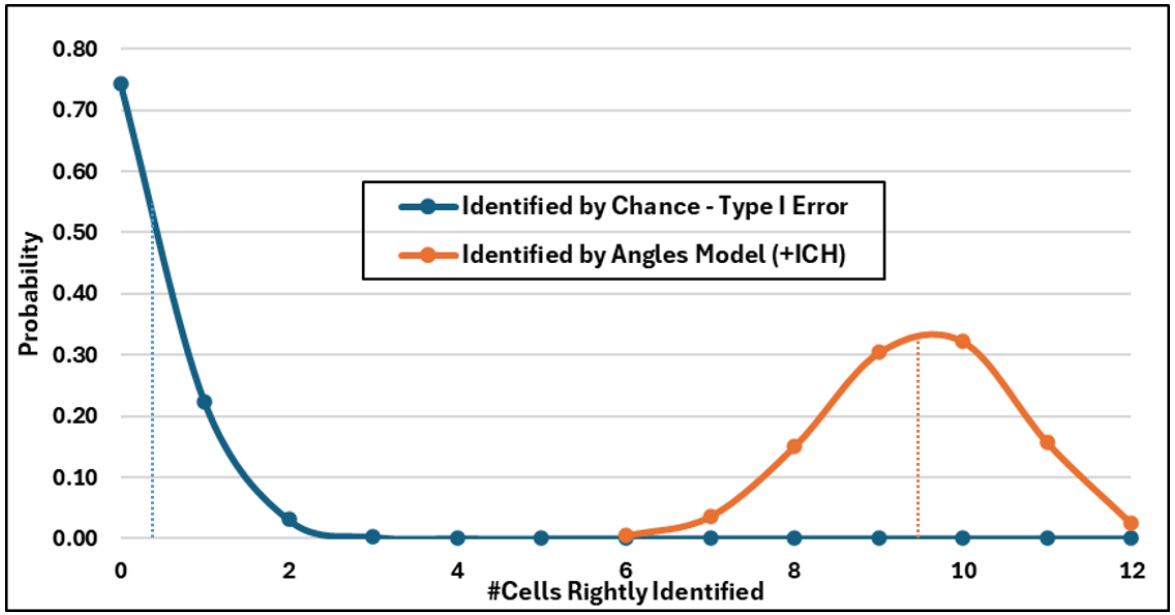
	I	C	E	G	L	S	M	J	D	Z
I										
C										
E	99.6%	99.5%								
G		99.8%								
L										
S			82.7%		76.9%					
M			100.0%							
J										
D			34.7%	68.9%	59.5%		51.4%	75.4%		
Z									98.5%	

	I	C	E	G	L	S	M	J	D	Z
I										
C	95.1%									
E										
G	96.2%		97.6%							
L	95.3%	95.8%	96.3%	97.0%						
S	97.1%	97.8%		97.4%						
M	97.7%	97.7%		96.6%	97.4%	98.3%				
J	95.6%	96.2%	96.8%	97.3%	97.5%	98.1%	97.9%			
D	97.2%	97.1%				98.0%				
Z	96.6%	97.6%	97.8%	98.1%	98.0%	98.0%	98.6%	99.0%		

V. Causal Discovery via DAG Recovery

- If we look only at the 12 cells with the directional relationships, the results remain compelling. Under the null of independence, the distribution of the number of right identifications becomes simply a binomial distribution with probability = Pr(A.), because there are no cells with 'no relationship' included. This is shown in Graph 23, and compared to the model's distribution of the number of right identifications for these 12 cells. The expected value of the former is approximately 0.29, and of the latter, 9.47. **This very material difference also is highly statistically significant**, with a p-value < 0.00001.* The angle model's %right value is 78.9% here (note that this is just recall, the percent of "positives" rightly identified), while that under the null is approximately 2.4%, which is Pr(A.).

Graph 23: Distribution of #Rightly Identified Directional Effects for 12 Pairs with Effects

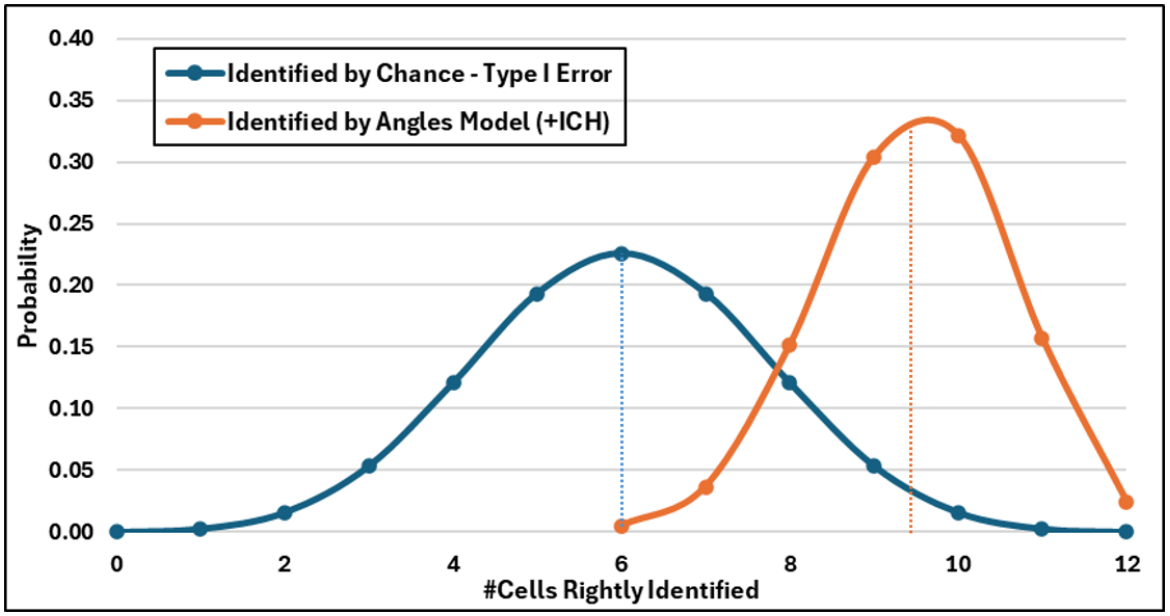


*** For this permutation test, sample sizes n1 = n2 = 12, and 100,000 simulations were run.**

V. Causal Discovery via DAG Recovery

- Even if we gave a researcher an unfair advantage over the angles-based model and he or she somehow KNEW that the 12 cells were associated with an effect (i.e. he or she had perfect priors), without knowing the direction of the effect, the distribution under the null would be a coin toss, with an expected value for the number of right selections of 6 out of the 12. Even under these conditions (see Graph 24), the model's effect is material and highly statistically significant, with a p-value of approximately 0.0001.* All of these results are summarized below in Table G.

Graph 24: Distribution of #Rightly Identified Directional Effects for 12 Pairs with Effects (Perfect Priors on which 12 Pairs Have SOME Effect, but not the Direction)



* For this permutation test, sample sizes n1 = n2 = 12, and 100,000 simulations were run.

V. Causal Discovery via DAG Recovery

TABLE G: Angles Distribution (ICH) DAG Recovery vs. No Model

All 45 Pairs, Fully Exploratory Analysis, No Priors

no model-type I error*		Angles-ICH		
# right	% right	# right	% right	p-value*
31.66	**70.4%**	**41.56**	**92.3%**	**<0.00001**

* Binomial vector: p12=adj_α*(1-adj_α), p33=p=[1 - adj_α]^2

* p-value: permutation test, nsims=100k, n1=n2=45

12 Pairs with Effects, Fully Exploratory Analysis, No Priors

no model-type I error**		Angles-ICH		
# right	% right	# right	% right	p-value**
0.29	**2.4%**	**9.47**	**78.9%**	**<0.00001**

** Binomial distribution, p12=adj_α*(1-adj_α)

** p-value: permutation test, nsims=100k, n1=n2=12

12 Pairs with Effects, Perfect Priors on Right Pairs but not Direction of Effect

no model-type I error~		Angles-ICH		
# right	% right	# right	% right	p-value~
6.00	**50.0%**	**9.47**	**78.9%**	**0.00068**

~ Binomial distribution, p12=0.5

~p-value: permutation test, nsims=100k, n1=n2=12

33 Pairs with No Effect, Fully Exploratory Analysis, No Priors

no model-type I error~~		Angles-ICH	
# right	% right	# right	% right
31.37	**95.1%**	**32.09**	**97.2%**

~~ Binomial distribution, p33=[1 - adj_α]^2

V. Causal Discovery via DAG Recovery

- This is a relatively narrow, preliminary study, but **the empirical results are promising, if not compelling**.
- **The study conditions are reasonably challenging**: all bivariate relationships are highly nonlinear, with heavy-tailed distributions and asymmetric error terms, **yet the angles-based model exhibits strong predictive power** even when sample sizes are relatively small (i.e. n=100, compared to other studies' use of n=1,000). These results (e.g. recall of 78.9%) compare favorably with other recent studies (see the recall of 76.7% for the winning DAG recovery method in an online contest with over 2,000 participants and 5,000 submissions, https://crunchdao.com/case-studies/adia-lab-causality).
- **The angle model's simultaneous, as opposed to sequential, estimation of edge calls, and the automatic enforcement of positive definiteness in the angles space, may very well contribute to greater coherence in DAG recovery**. Results will vary depending on the (directional) dependence measure used, but this is a strength rather than a weakness of this approach: it is adaptable to different and varying data conditions, under which some measures will perform better than others. So **having the flexibility to verify data conditions ex ante could lead to notable power gains**.
- This all will be explicitly addressed in a more extensive follow-up empirical study that will include a wide range of sample sizes, many different data types with varying signal-to-noise ratios, and results from competitor algorithms designed for identifying causal structure.
- But these preliminary results, for a completely original methodology, are very encouraging.

V. Causal Discovery via DAG Recovery

- Of course, this dive into (directional) causality and DAG recovery all begs the bigger and valid question of whether DAGs can be used reliably within "self-referencing open systems like capital markets" to begin with (Polakow et al., 2023).

- Importantly, many express strong caution, based on recent and rigorous research, regarding the reliance on DAGs in this setting (see de Lara, 2023; Gong et al., 2024).*

- I propose only that the angles-based model appears to be able to play an effective role here in identifying causal structure and accurately recovering DAGs if the answer to this question turns out to be "yes" or "under some conditions."**

*** From Polakow et al. (2023): "The clarion call for causal reduction in the study of capital markets is intensifying. However, in self-referencing and open systems such as capital markets, the idea of unidirectional causation (if applicable) may be limiting at best, and unstable or fallacious at worst." From Gong et al. (2024): "… potential outcomes (PO) and structural causal models (SCMs) stand as the predominant frameworks. However, these frameworks face notable challenges in practically modeling counterfactuals … we identify an inherent model capacity limitation, termed as the 'degenerative counterfactual problem', emerging from the consistency rule that is the cornerstone of both frameworks." And from De Lara (2024): "Most of the literature on causality considers the structural framework of Pearl and the potential-outcomes framework of Neyman and Rubin to be formally equivalent, and therefore interchangeably uses the do-notation and the potential-outcome subscript notation to write counterfactual outcomes. In this paper, we … prove that structural counterfactual outcomes and potential outcomes do not coincide in general – not even in law." See Opdyke (2024b) for a more complete review of this literature.**

**** Note here the recent publication of Reisach et al., 2025, which proposes a formalization via composite causal variables to explicitly incorporate time into DAGs. This is complementary to the earlier work of Rodriguez Dominguez (2023, 2025a).**

VI. Inference and Sampling Even Under Restricted Scenarios

2. Fully General Case: Any Positive Definite Dependence Measure, Any Values, Any Data

A NEW METHOD FOR SCENARIOS, FULLY FLEXIBLE WITHIN THE ALL-PAIRWISE FRAMEWORK

- Note that if we are conducting the causal discovery analysis as in the previous section above, but believe that specific bivariate relationships within the 'system' of variables will remain unaffected by the scenario we are testing, we cannot simply exclude those factors, as they affect other factors that remain relevant to the scenario.
- In other words, freezing the values of bivariate pairs is not the same as freezing the effects of one factor on all others.
- Addressing this need for fully flexible scenarios requires devising a method that works on the entire matrix / portfolio of factors and perturbs the angles space, but **freezes the individual cells** unaffected by the scenario **while still preserving the inferential validity of the rest of the matrix**. Stated differently, we must obtain the finite sample distribution of the scenario-restricted matrix.
- Nothing in the extant literature provides this.
- To be clear, **even though we are making inferences in angles space, what we are ultimately concerned about here is freezing the unaffected cell values of the actual dependence matrix**. The re-sorting method presented below achieves this, for a large number of the possible combinations of 'frozen' cells within the framework of the all-pairwise matrix. It provides unprecedented scenario flexibility when it comes to inference related to dependence structures.

VI. Inference and Sampling Even Under Restricted Scenarios

2. Fully General Case: Any Positive Definite Dependence Measure, Any Values, Any Data

A NEW METHOD FOR SCENARIOS, FULLY FLEXIBLE WITHIN THE ALL-PAIRWISE FRAMEWORK

- Several characteristics, properly taken together, provide the angles approach with the ability to 'freeze' selected combinations of scenario-specified cells within the framework of the all-pairwise matrix, while **maintaining both inferential and representative sampling validity for the remaining (non-frozen) matrix**.
  - i. Independence of the angles distributions
  - ii. Invariance of the distributions of individual cells, and the entire matrix, to the orderings of the rows and columns of the matrix
  - iii. The mechanics of matrix multiplication
  - iv. The granular, cell-level geometry of the angles approach

- First, recall that i. the angles distributions are independent, and ii. that the distributions of both the individual cells of the matrix, and that of the entire matrix, remain invariant to the orderings of the rows and columns of the matrix (see Pourahmadi & Wang, 2015, and Lewandowski et al., 2009). Based on i. and ii., we can exploit iii., the simple mechanics of matrix multiplication, so that only selected cells of the matrix are affected, and the rest frozen, as required for a given scenario.

- To explain iii., I focus only on the lower triangle of the correlation matrices below in Graphs 15-17, since the upper triangle is just its reflection due to symmetry.

VI. Inference and Sampling Even Under Restricted Scenarios

2. Fully General Case: Any Positive Definite Dependence Measure, Any Values, Any Data

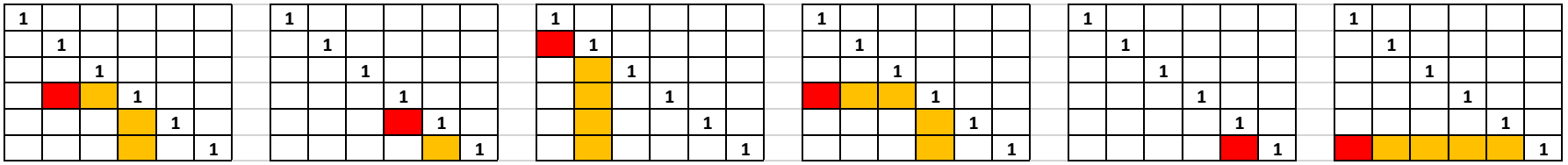
- Note again that using angles here, **we never perturb the correlation/dependence measure values directly. We must always convert correlations to angles, perturb the angle values using the fitted kernels, and then if needed, translate back to correlation values**.

- In doing so, when multiplying the Cholesky factor by its transpose, $R = BB^T$, or equivalently,

$$r_{i,j} = \cos(\theta_{i,1})\cos(\theta_{j,1}) + \sum_{k=2}^{i-1}\left[\cos(\theta_{i,k})\cos(\theta_{j,k})\prod_{l=1}^{k-1}\sin(\theta_{i,l})\sin(\theta_{j,l})\right] + \cos(\theta_{j,i})\prod_{l=1}^{i-1}\sin(\theta_{i,l})\sin(\theta_{j,l}) \quad \text{for } 1 \le i < j \le n$$

changing an arbitrary angle in *B* only will change any correlations that are to its right in the same row, and under the diagonal in the corresponding column.* Several examples of this are shown below.

Graph 15: Correlation Cells Affected by Changing a Specific Angle Cell



*** Note that not all of these (orange) correlation cells will necessarily change if values of zero are involved, but none OTHER than these (orange) correlation cells CAN change when only the red angle cell changes.**

VI. Inference and Sampling Even Under Restricted Scenarios

2. Fully General Case: Any Positive Definite Dependence Measure, Any Values, Any Data

- This means that we can simply reorder the matrix so that the targeted cells we want to vary all end up in the rightmost triangle of the lower triangle of the matrix, according to the fill order in Graph 16 below. If we only change in matrix *B* the angle values of cells 1, 2, and 3 below, no other cells in the correlation matrix *R* will be affected, simply by virtue of the mechanics of matrix multiplication from $R = BB^T$. Due to i. and ii., **this re-sorting does not affect the distributions of the specific cells nor that of the entire matrix**.

Graph 16: 'Fill Order' for Resorting the Correlation Matrix to Change Only Specified Cells

Rightmost Triangle Fill Order

11					
12	7				
13	8	4			
14	9	5	2		
15	10	6	3	1	

VI. Inference and Sampling Even Under Restricted Scenarios

2. Fully General Case: Any Positive Definite Dependence Measure, Any Values, Any Data

Graph 17: Matrix Resorting Using the 'Fill Order' to Change Only Specified Correlation Cells

Determine Targeted Change Cells

1,2					
1,3	2,3				
		4,3			

Reorder Rows/Cols to Fill Rightmost Triangle with Targets According to Fill Order

11					
12	7				
13	8	4			
14	9	5	2		
15	10	6	3	1	

Changes in Corresponding Angles Cells ONLY change Same in Resorted Matrix

11					
12	7				
13	8	4,3			
14	9	5	2,3		
15	10	6	1,3	1,2	

- We see in Graph 17 that reordering the correlation matrix so that rows 1-6 are now 6-1 and columns 1-6 are now 6-1, means that the original (targeted) cells 1,2 and 1,3 and 2,3 and 4,3 are now in the rightmost triangle of the lower triangular matrix, in the fill order shown in Graph 16, and changes to the corresponding cells in the angles matrix *B* will only change these same cells, after $R = BB^{T}$, in the resulting correlation matrix.
- Note that the targeted cells (green) do not have to be contiguous. Once they are re-ordered into the lower right triangle, **changes in these (orange) cells in the angles matrix will ONLY change these corresponding (orange) cells in the correlation matrix!**
- Readers are encouraged to test this in the fully interactive spreadsheet (rows 122-125) found at http://www.datamineit.com/DMI_publications.htm

VI. Inference and Sampling Even Under Restricted Scenarios

2. Fully General Case: Any Positive Definite Dependence Measure, Any Values, Any Data

- Changes to targeted cells WILL change other cells in the targeted submatrix (due to $R = BB^T$). HOWEVER, the ordering of the submatrix matters here, and can be exploited. For example, if we want to change the 4 cells above, but subsequently want to perform 'what if' analyses on only one of those cells (e.g. cell 1,2) without changing the other three, we order the original correlation matrix to place that cell as the 'first' in the lower triangle of the *B* matrix. Then, subsequent changes to it will not affect the other (orange) cells. **This 'rightmost' change rule is nested / hierarchical!**
- Note that in this example matrix, there are [p(p-1)/2]! = 15! = 1,307,674,368,000 possible permutations of the cells for (ordered) selection for a scenario, but only 6! = 720 ways to sort the rows and columns within the framework of the all-pairwise matrix. Freely sorting the cells per the former obviously breaks the structure of the all-pairwise matrix, so every conceivable scenario corresponding to every possible grouping of cells is not attainable (this is the price of relying on the all-pairwise matrix structure). So for some scenarios, we will be forced to include in the lower right triangle one or a few cells that do not belong. But given the size of the matrices used in practice, these are relatively rare occurrences, and when they do arise, the effects of these unwanted inclusions can be tested for materiality via 'what if' analyses.
- Finally, note that this re-sorting approach would not be possible if its geometry rested on more complex, multivariate dependence measures, as opposed to the bivariate all-pairwise matrix.
- BOTTOM LINE: **the scenario flexibility provided by this re-sorting approach, while simultaneously preserving inferential (and sampling) validity, eclipses any other method in the extant literature**.

VI. Inference and Sampling Even Under Restricted Scenarios

2. Fully General Case: Any Positive Definite Dependence Measure, Any Values, Any Data

- One question remains: **what values do we use for those cells that are 'frozen'?** As alluded to above when dealing with asymmetry for angles-based two-sided p-values, **we use the mean correlation matrix**. Conduct the N (=10k) simulations, as before, but for the 'frozen' cells, always insert the mean angle values that correspond to the mean correlation matrix. This effectively 'freezes' these cells to their estimated values, while preserving the positive definiteness of the matrix (inserting mean cell values will place the matrix deeper into the convex cone, so this should never, as an empirical matter, render the matrix non-positive definite).
- Extensive simulations, not presented herein, empirically validate this re-sorting method with respect to the above-described mean substitution, although the cell-freezing mechanism is fully deterministic.

VII. Beyond Distance – A Measure of Generalized Entropy

- Yet another application of the angles-space relates to measuring 'distance' vis-à-vis the all-pairwise matrix of positive definite dependence measures, and bringing probabilistic meaning to such distance measurements.

- It turns out that the (two-sided) cell-level p-values in the angles space, as defined above, can be used to construct **a competitor to commonly used distance metrics, such as norms, and it has a number of advantages over them in this setting**.

- Some commonly used norms for measuring correlation 'distances' include the Taxi, Frobenius/Euclidean, and Chebyshev norms (collectively, the Minkowski norm), shown below

$$\|x\| = \left(\sum_{i=1}^{d} |x_i|^m \right)^{1/m}$$

where x is a distance from a presumed or baseline correlation value, d=number of observations, and m=1, 2, and ∞ correspond to the Taxi, Frobenius/Euclidean, and Chebyshev norms, respectively.

VII. Beyond Distance – A Measure of Generalized Entropy

- All of these norms measure absolute distance from a presumed or baseline correlation/dependence measure value. But the range of all relevant and widely used dependence measures is bounded, either from –1 to 1 or 0 to 1, and **both the relative impact and *meaning* of a given distance at the boundaries are not the same as those in the middle of the range**.

- In other words, a shift of 0.02 from an original or presumed correlation/dependence value of, say, 0.97, means something very different than the same shift from 0.27. **Angle-space p-values attribute probabilistic MEANING** to these two different cases, while a norm would treat them identically, even though they very likely indicate what are very different events, of very different relative magnitudes, with potentially very different consequences.

VII. Beyond Distance – A Measure of Generalized Entropy

- Therefore, a natural, PROBABILISTIC distance measure, based directly on angle-space p-values, is the natural log of the product of the p-values, dubbed 'LNP' below:

$$\text{"LNP"} = \ln\left(\prod_{i=1}^{p(p-1)/2} p\text{-}value_i\right) =$$

$$= \sum_{i=1}^{p(p-1)/2} \ln\left[p\text{-}value_i\right] \text{ where angles-based } p\text{-}value_i \text{ is two-sided}$$

- Using a Pearson's correlation matrix under the (Gaussian) identity matrix, **LNP shows a very strong correspondence with the entropy of the correlation matrix**, defined by Felippe et al. (2021 and 2023) as below:

$$\text{Entropy} = Ent\left(R/p\right) = -\sum_{j=1}^{p} \lambda_j \ln\left(\lambda_j\right)$$

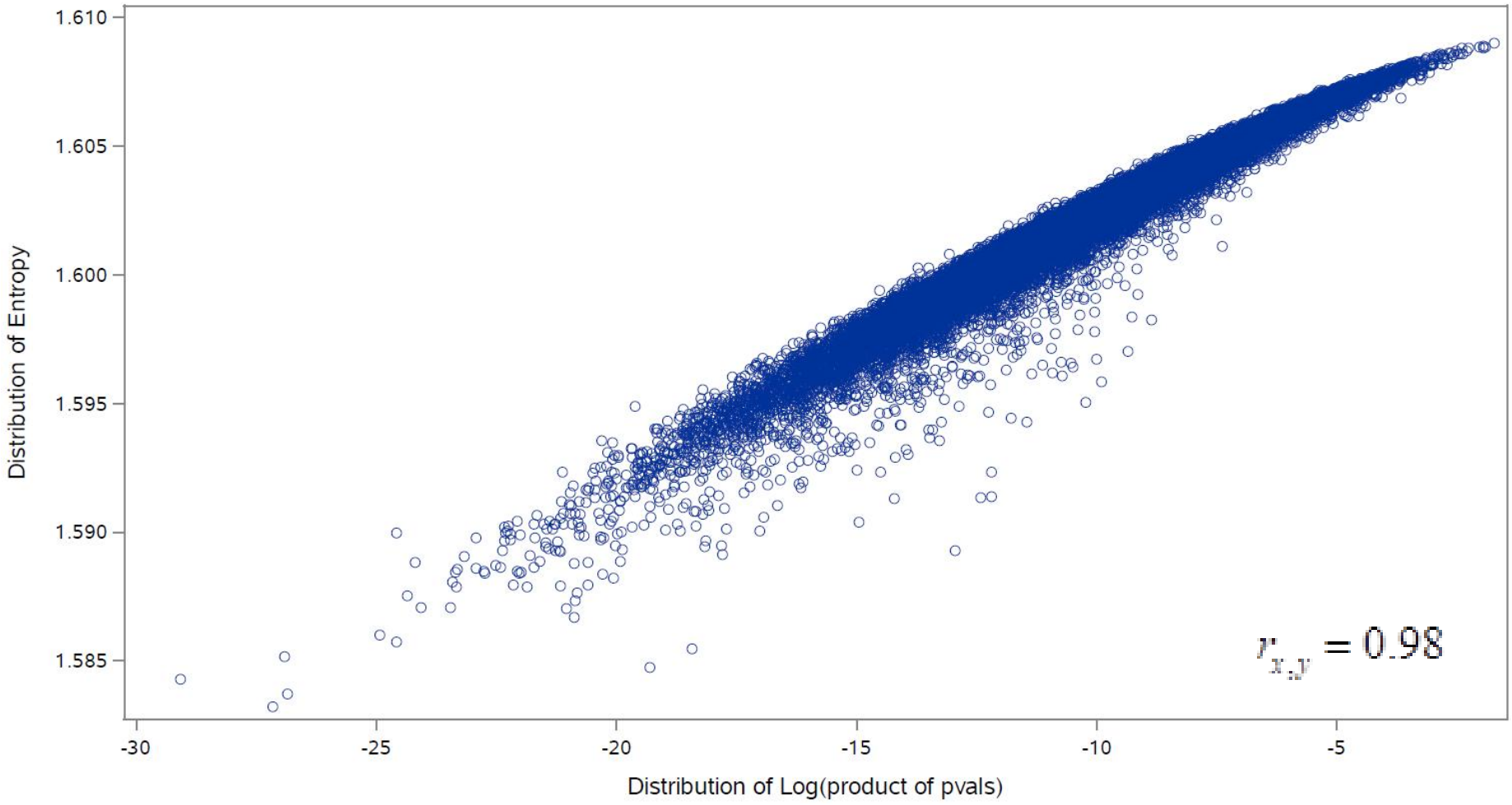
where *R* is the sample correlation matrix and λ_j are the p eigenvalues of the correlation matrix after it is scaled by its dimension via *R*/*p*.

- Importantly, this entropy result also is valid for ANY positive definite measure of dependence, not just Pearson's.

- Graph 18 below compares LNP vs entropy of Kendall's Tau matrix in 10,000 simulations (with n=126 for half a year of daily returns) under the Gaussian identity matrix. The Pearson's correlation between them is (0.98). I obtain almost identical results when LNP is compared to the entropy of Pearson's, Spearman's, and the symmetric version of Chatterjee's correlation.

VII. Beyond Distance – A Measure of Generalized Entropy

GRAPH 18: Identify Matrix Simulations for Kendall's Tau – LNP vs. Correlation Matrix Entropy



VII. Beyond Distance – A Measure of Generalized Entropy

- It is important to note, however, that **entropy here is limited** to being calculated relative to the case of independence, which for many dependence measures corresponds only with the identity matrix.*
- In contrast, **LNP can be calculated, and retains its meaning, in all cases, based on ANY values of the dependence matrix, not just the case of independence**.
- Yet the correspondence of LNP to entropy under this specific case speaks to LNP's natural interpretation as a meaningful measure of distance or deviation or disorder (depending on your interpretation), and one that also is **more flexible and granular than entropy as it is measured cell-by-cell, p(p-1)/2 times, as opposed to only p times for p eigenvalues**.
- As such, LNP might be considered a type of 'generalized entropy' relative to ANY baseline values of ANY positive definite dependence measure, including as a special case those values corresponding with perfect (in)dependence.
- Note also that because LNP is a monotonic transformation, the relative rankings of different dependence measure matrices will be preserved, which is useful when comparing LNPs based on simulations to an estimated/observed LNP.
- Also, unlike norms, LNP's scale is not dependent on sample size, although it is dependent on the dimension of the matrix, p.

*** Recall, of course, that a zero value for Pearson's or Kendall's or Spearman's does not imply independence, but independence does imply a zero value for these measures (the exception being Pearson's under Gaussian data, for which a zero value *does* indicate independence).**

VII. Beyond Distance – A Measure of Generalized Entropy

- Such entropy-related measures certainly are relevant in this setting as entropy has been used increasingly in the literature to measure, monitor, and analyze financial markets (see Meucci, 2010b, Almog and Shmueli, 2019, Chakraborti et al., 2020, and Vorobets, 2024, 2025, for several examples).

- So the use of **LNP** here warrants further investigation as a matrix-level measure of distance that, unlike widely used distance measures such as norms, **has a meaningful probabilistic foundation**.

- **It applies not only beyond the independence case generally, but also to ALL positive definite measures of dependence regardless of their values**.

- **LNP is a monotonic transformation that preserves relative rankings, and unlike norms, its scale is not a function of sample size.**

- These are intriguing results with possibly far-reaching implications.

VIII. Additional Applications / Further Research

- Additional research can potentially increase the utility and breadth of the angles-based approaches presented herein for causal discovery, inference, sampling, and distance measurement.

- Analytic Angles Distributions, if not for the General Case, then additional Special Cases: I provide above the fully analytic derivation of the angles distribution under the Gaussian identity matrix: a foundational, but admittedly narrow special case. Although the angle approach's general solution remains 'runtime reasonable' given its generality and objectives, expanding the range of conditions for a fully analytic solution would dramatically speed up it's implementation. Deriving an "all cases" analytic solution currently appears to be a nontrivial problem, but even providing this under additional specific cases would be very useful and directly useable in it's application.

- A Second, Follow-up Empirical Study Using the Angles Approach for DAG Recovery: This is currently under way, and will include a wide range of sample sizes, and many different data types with varying signal-to-noise ratios. Specifically, the potential advantages of the unique approach provided by the angles space, including

  i. increased coherence in edge calls (due to automatic enforcement of positive definiteness, as well as simultaneous rather than sequential edge calling), and

  ii. greater power under wide-ranging data conditions due to flexibility in its use of any positive definite dependence measure

  will be tested against competitor algorithms.

IX. Conclusions

- This presentation, which includes results from an accompanying monograph, utilizes the angle space of positive definite dependence measures to derive original methods for 1. causal discovery; 2. inference for and sampling of these dependence measures; and 3. distance measurement using these measures.
- Positive definite dependence measures arguably span all those in widespread usage, so this approach maintains a broad and relevant range of application.
- **For causal discovery**, the angles space provides an original mechanism for DAG Recovery with promising, if not compelling initial empirical results. Compared to competitors, its automatic enforcement of positive definiteness and simultaneous rather than sequential edge calling could provide increased coherence, and its flexibility to use any directional positive definite dependence measure could lead to notably increased power. A follow-up study will test these comparisons.
- **For inference and sampling**, the angles space provides a new, fully analytic derivation solving a foundational inference problem for dependence structure, as well as sampling for the same 30% faster than the current fastest methods in the extant literature.
- **For distance measurement**, the angles space provides a new competitor to commonly used distance metrics, such as norms: a “generalized entropy” that provides probabilistic meaning to the distance measured by positive definite dependence measures. It is a monotonic transformation whose scale does not depend on sample size; its range of application is broader than matrix entropy, and its measurement is more granular, on the right (same) level of aggregation for (as) the all-pairwise dependence measure matrix.
- The angles space provides numerous other advantages in these settings; it’s already wide breadth of useful application undoubtedly will continue to expand rapidly in the relevant literature.

Author Contact

Contact Information:

JD Opdyke
Chief Analytics Officer, DataMineit, LLC
JDOpdyke@DataMineit.com
JDOpdyke@gmail.com

X. Appendix: Assumptions for Valid Causal Interpretation

A. Below I define the assumptions under which the directional associations identified by NAbC can be interpreted causally, and the causal meaning of edges inferred by NAbC in a DAG.

Preliminaries: Structural Causal Model (SCM) and do-Semantics (see Pearl, 2009)

Let $V = \{X_1, \ldots, X_p\}$ be observed variables generated by an acyclic Structural Causal Model (SCM):

$$X_j = f_j\left(X_{Pa(j)}, U_j\right) \text{ for } j = 1, \ldots, p$$

Here Pa(j) are the parents of X_j in a DAG "G," $U_1, \ldots, U_p$ are jointly independent exogenous noises, and each $f_j(\ )$ is a (possibly nonlinear, nonparametric) measurable function. The graph G encodes the causal Markov property and admits do-interventions: for any variable X, the operation do(X:=x) replaces its structural equation with the constant x, severing incoming edges into X.

Causal effect: X has a (possibly context-dependent) causal effect on Y if and only if there exist x ≠ x' such that P(Y | do(X=x)) ≠ P(Y | do(X=x')).

Assumptions for Causal Interpretation:

(A1) Acyclicity: The true causal graph over V is a DAG.

(A2) Causal Sufficiency (no latent confounders among V): The exogenous noises U_j are mutually independent; any common causes of variables in V are included in V. Relaxation: if (A2) is uncertain, interpret edges as ancestral (existence of a directed path) rather than necessarily adjacent (direct).

(A3) Markov + (qualified) faithfulness: The observed joint distribution is Markov with respect to G and violations of direction-revealing asymmetries are non-degenerate (i.e. pathological cancellations are excluded).

(A4) i.i.d. sampling (or weak stationarity for time-series): Samples are independent draws from the same distribution (future research plans to apply NAbC for time-series extensions, to innovations or appropriate lags).

(A5) Measurement reliability: Variables are measured without systematic differential error that inverts directional asymmetries.

(A6) Conditioning set for direct edges: When inferring adjacency (direct edges), directional dependence is evaluated conditionally on V\{X,Y}. If only marginal directional dependence is used, interpret edge calls as directed reachability (existence of some directed path).

(A7) Monotone link from mechanism to asymmetry: For the chosen directional measure *M* (in the empirical study above, this is ICH – the improved Chatterjee's correlation of Xia et al., 2024), if X ∈ Pa(Y)

X. Appendix: Assumptions for Valid Causal Interpretation

and Y ∉ Pa(X), then asymptotically E[M(X→Y)] > E[M(Y→X)], with separation increasing in sample size n under the SCM.

Note that this assumption plays a fundamental role in linking the structural causal model to observable directional asymmetries. It requires that the directional dependence functional M increase in expectation in the true causal direction. This monotone link between mechanism and measurable asymmetry is the critical identifiability condition for NAbC-based DAG recovery. In practice, validation of (A7) proceeds through a simulation across a range of nonlinear, non-Gaussian, and/or heteroskedastic settings, ensuring that E[M(X→Y)] > E[M(Y→X)] consistently emerges under the data-generating process.

Among the seven assumptions, only (A7) must be verified empirically.

Edges in NAbC:

Let M be a directional dependence functional and $\hat{M}$ its sample estimate. NAbC maps the full matrix of $\hat{M}$ values to angles, performs finite-sample joint inference under positive-definiteness, and returns p-values for directed hypotheses.

Adjacency claim (direct edge): Under (A1)–(A7) and using conditional directional dependence M(X→Y | V\{X,Y}): declare a directed edge X→Y if statistically significant(X→Y) and NOT statistically significant(Y→X) at level α (the one-direction-only rule). This is interpreted as X ∈ Pa(Y).

Ancestral/path claim (directed reachability): If only marginal M is used, interpret the same rule as: there exists a directed path X ⇝ Y (not necessarily adjacent); if (A2) is violated, reduce the claim to directional dependence without causal interpretation.

No-edge claim: If neither direction is statistically significant, refrain from a causal claim.

Cycle handling: If the one-direction-only rule yields a 2-cycle, pruning enforces acyclicality by selecting the edge/direction that has a larger p-value and replacing its dependence measure value with a value just below the α threshold of statistical significance. The same is done when longer cycles are encountered. The specific algorithm for the latter case that was implemented in the above empirical study used Kahn's (1962) topological sorting algorithm to identify cycles, and then i. selected those nodes/variables with the fewest edges; ii. then among those, selected the one with the largest sum of p-values across its related edges (i.e. the weakest cumulative signals/edges); and iii. finally, selected the particular edge from among those that had the largest p-value (weakest signal/edge). This process was repeated until no cycles remained. In the above empirical study, no pruned matrices became non-positive definite, which is not surprising given that the pruning adjustments place the matrix further r within the positive definite cone (still, positive definiteness always was empirically verified). But even if positive definiteness had been occasionally violated, non-linear constrained optimization is relatively straightforward to implement here. While this pruning algorithm is not multivariate optimized, it is fast and straightforward, and was effective in the empirical study herein. A follow-up study will explore optimizing this approach, possibly by minimizing an aggregate angle-space loss function under positive definite constraints.

X. Appendix: Assumptions for Valid Causal Interpretation

Under (A1)–(A7), the one-direction-only NAbC decision rule using conditional *M* is pointwise sound for adjacency: if the rule declares X→Y, then X ∈ Pa(Y) with probability → 1 as n → ∞.

By Markov + faithfulness (A3) and SCM acyclicity (A1), conditioning on V\{X,Y} blocks all non-causal paths between X and Y. If X ∈ Pa(Y), then by (A7) the population asymmetry satisfies *M*(X→Y) > *M*(Y→X); NAbC's finite-sample tests are consistent in n (angles-based inference), so the probability of correctly calling X→Y tends to 1. Conversely, if no directed edge X→Y exists, all back-door and collider paths are blocked by the conditioning set, implying symmetric (null) directional dependence; the rule refrains from a causal call with probability → 1.

Note that when *M* is conditional on other variables, an edge X→Y indicates a direct causal relationship, but when *M* is marginal, the same edge indicates only a causal pathway or reachability, not necessarily a direct connection. In the empirical study in Section 8. above, no conditioning was used, and consequently, the directed edges recovered therein should be interpreted as reachability rather than as direct causal effects.

Finally, note again that NAbC, as defined herein, identifies causal structure, but does not estimate effect sizes or counterfactual outcomes.

B. Below I place NAbC within the Causal Inference Literature

Proposition:

Let $V = \{X_1, \ldots, X_p\}$ be generated by acyclic structural equations $X_j = f_j\left(X_{Pa(j)}, U_j\right)$ with mutually independent exogenous noises U_j. Under the causal Markov property and (qualified) faithfulness, the observed distribution is Markov to the causal DAG "G." I adopt these semantics and make the following explicit: (i) acyclicity; (ii) causal sufficiency (or an ancestral interpretation if violated); (iii) i.i.d./weak-stationarity; (iv) measurement reliability. Directional inference in NAbC uses a directional dependence functional *M*; edges are interpreted as direct parents when *M* is evaluated conditionally on V\{X,Y}, and as directed reachability when evaluated marginally.

This makes NAbC comparable with established approaches in the causal literature (see Zanga et al., 2025). Some of these include constraint-based methods (e.g. PC-Stable/FCI, see Colombo and Maathuis, 2014), score-based methods (e.g. FGES, NOTEARS, see Ramsey et al., 2017, and Zheng et al., 2018), and functional/ANM methods (LiNGAM/DirectLiNGAM, see Hyvärinen et al., 2010, and Shimizu et al., 2011) among others. Unlike those examples, however, NAbC takes a completely original approach by performing finite-sample inference with positive-definite constraints at the matrix level, which may very well improve coherence of edge calls across the graph, especially compared to those algorithms that identify causal edges sequentially. This will be explored in a more extensive, future empirical study pitting NAbC against many of the above-mentioned competitors.

XI. References

Abul-Magd, A., Akemann, G., and Vivo, P., (2009), "Superstatistical Generalizations of Wishart-Laguerre Ensembles of Random Matrices," *Journal of Physics A Mathematical and Theoretical*, 42(17):175207.

Adams, R., Pennington, J., Johnson, M., Smith, J, Ovadia, Y., Patton, B., Saunderson, J., (2018), "Estimating the Spectral Density of Large Implicit Matrices" https://arxiv.org/abs/1802.03451.

Adegoke, N., Ajadi, J., Mukherjee, A., and Abbasi, S., (2022), "Nonparametric Multivariate Covariance Chart for Monitoring Individual Observations," *Computers & Industrial Engineering*, Vol 167.

Afriat, S.,(1957), "Orthogonal and oblique projectors and the characterization of pairs of vector spaces," *Mathematical Proceedings of the Cambridge Philosophical Society*, 53 (4): 800–816.

AghaKouchak, A., Easterling, D.,Hsu, K., Schubert, S., and Sorooshian, S., eds, (2013), Extremes in a Changing Climate: Detection, Analysis and Uncertainty, Ch. 6: Methods of Tail Dependence Estimation (pp.163-179), Springer Nature, Part of the book series: Water Science and Technology Library (WSTL, volume 65).

Ajadi, j., Wong, A., Mahmood, T., and Hung, K., (2021), "A new multivariate CUSUM chart for monitoring of covariance matrix with individual observations under estimated parameter," *Quality and Reliability Engineering International*, 38(2), 834-847.

Akemann, G., Fischmann, J., and Vivo, P., (2009), "Universal Correlations and Power-Law Tails in Financial Covariance Matrices," https://arxiv.org/abs/0906.5249.

Allena, D., and McAleerb, M., (2018), "'Generalized Measures of Correlation for Asymmetry, Nonlinearity, and Beyond': Comment," ICAE – Instituto Complutense de Análisis Económico, Working Paper 1823.

Almog, A., and Shmueli, E., (2019), "Structural Entropy: Monitoring Correlation-Based Networks over time With Application to Financial Markets," *Scientific Reports*, 9:10832.

Alonso, M., (2025), "Return-Adjusted Hierarchical Risk Parity and Schur Portfolios: A Comprehensive Theoretical and Empirical Study," working paper, https://ssrn.com/abstract=5370624.

Alpay, D., and Mayats-Alpay, L., (2023), "Similarity Metrics, Metrics, and Conditionally Negative Definite Functions," arXiv:2307.10446v1 [math.FA].

Andu, Y., Lee, M., and Algamal, Z., (2021), "Adaptive Elastic Net with Distance Correlation on the Group Effect and Robust of High Dimensional Stock Market Price," *Sains Malaysiana*, 50(9), 2755-2764.

Archakov, I. and Hansen, P., (2021), "A New Parametrization of Correlation Matrices," *Econometrica*, 89(4), 1699-1715.

Aznar, D., (2023), "Portfolio Management: A Deep Distributional RL Approach," University of Barcelona, Thesis.

Babić, S., Ley, C., Ricci, L., and Veredas, D., (2023), "TailCoR: A new and simple metric for tail correlations that disentangles the linear and nonlinear dependencies that cause extreme comovements." *PLoS ONE*, 18(1): e0278599

Bank for International Settlements (BIS), (2011a), Basel Committee on Banking Supervision, Working Paper 19, (1/31/11), "Messages from the academic literature on risk measurement for the trading book."

Bank for International Settlements (BIS), (2011b), Basel Committee on Banking Supervision, "Operational Risk – Supervisory Guidelines for the Advanced Measurement Approaches," June, 2011.

Barber, R., and Kolar, M., (2018), "ROCKET: Robust Confidence Intervals via Kendall's Tau for Transelliptical Graphical Models," *The Annals of Statistics*, 46(6B), 3422-3450.

Benjamini, Y., and Hochberg, Y., (1995), "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing," *Journal of the Royal Statistical Society*, Series B. 57 (1): 289–300.

Besson, O., (2025), "Covariance Matrix Estimation in the Singular Case Using Regularized Cholesky Factor," arXiv:2505.16302v1 [math.ST].

Björck, Å. and Golub, G., (1973), "Numerical Methods for Computing Angles Between Linear Subspaces," *Mathematics of Computation*, 27 (123): 579–594.

Blakesley, R., Mazumdar, S., Dew, M., Houck, P., Tang, G., Reynolds, C., and Butters, M., (2009), "Comparisons of Methods for Multiple Hypothesis Testing in Neuropsychological Research," *Neuropsychology*, 23(2), 255-264.

Blömbaum, P., Janzing, D., Washio, T., Shimizu, S., and Schölkopf, B., (2019), "Analysis of Cause-Effect Inference by Comparing Regression Errors," *Peer Journal of Computational Science*, 5:e169. doi: 10.7717/peerj-cs.169.

Blomqvist, N., (1950) "On a Measure of Dependence between Two Random Variables", *Annals of Mathematical Statistics*, 21(4): 593-600.

Bonferroni, C. E., (1936), "Teoria statistica delle classi e calcolo delle probabilità," *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze*.

Bongiorno, C., Challet, D., and Loeper, G., (2023), "Filtering time-dependent covariance matrices using time-independent eigenvalues," *Journal of Statistical Mechanics: Theory and Experiment*, Vol. 2023.

Bongiorno, C., and Challet, D., (2023a), "Covariance Matrix Filtering and Portfolio Optimization: The Average Oracle vs Non-linear Shrinkage and All the Variants of DCC-NLS", arXiv:2309.17219v1 [q-fin.ST].

Bongiorno, C., and Challet, D., (2023b), "Non-linear Shrinkage of the Price Return Covariance Matrix is Far from Optimal for Portfolio Optimization," *Finance Research Letters*, Vol. 52, 103383.

Bouchaud, J, & Potters, M., (2015), "Financial applications of random matrix theory: a short review," The Oxford Handbook of Random Matrix Theory, Eds G. Akemann, J. Baik, P. Di Francesco.

Bouchaud, J., (2021), "Radical Complexity," *Entropy*, Vol. 23.

Bours, M., and & Steland, A., (2020), "Large-sample approximations and change testing for high-dimensional covariance matrices of multivariate linear time series and factor models," *Scandinavian Journal of Statistics*, 48(2), 610-654.

Bulut, H., (2025), "A Novel Robust Test to Compare Covariance Matrices in High-Dimensional Data," *Axioms*, 14, 247.

Burda, Z., Jurkiewicz, J., Nowak, M., Papp, G., and Zahed, I., (2004), "Free Levy Matrices and Financial Correlations," *Physica A: Statistical Mechanics and its Applications*.

Burda, Z., Gorlich, A., and Waclaw, B., (2006), "Spectral Properties of empirical covariance matrices for data with power-law tails," *Phys. Rev., E 74*, 041129.

Burda, Z., Jaroz, A., Jurkiewicz, J., Nowak, M., Papp, G., and Zahed, I., (2011), "Applying Free Random Variables to Random Matrix Analysis of Financial Data Part I: A Gaussian Case," *Quantitative Finance*, Volume 11, Issue 7, 1103-1124.

Burda, Z., and Jarosz, A., (2022), "Cleaning large-dimensional covariance matrices for correlated samples," *Phys. Rev. E*, 105, 034136.

Cai, Y., Li, X., Sun, M., and Li, P., (2025), "Recovering Linear Causal Models with Latent Variables via Cholesky Factorization of Covariance Matrix," In: Le Thi, H.A., Le, H.M., Nguyen, Q.T. (eds) Advances in Data Science and Optimization of Complex Systems, ICAMCS 2024, Lecture Notes in Networks and Systems, Vol. 1311, Springer, Cham.

Cardin, M., (2009), Multivariate Measures of Positive Definiteness," *International Journal of Contemporary Mathematical Sciences*, 4(4), 191-200.

Carrara, C., Zambon, L., Azzimonti, D., and Corani, G., (2025), "A Novel Shrinkage Estimator of the Covariance Matrix for Hierarchical Time Series," in di Bella, E., Gioia, V., Lagazio, C., Zaccarin, S. (eds), Statistics for Innovation I, SIS 2025, Italian Statistical Society Series on Advances in Statistics, Springer, Cham.

Casa, A., and Cappozzo, A., (2025), "Penalized Model-Based Clustering for Covariance Matrices," in di Bella, E., Gioia, V., Lagazio, C., Zaccarin, S. (eds), Statistics for Innovation I, SIS 2025, Italian Statistical Society Series on Advances in Statistics, Springer, Cham.

Centofanti, F., Hubert, M., and Rousseeuw, P., (2025), "Cellwise and Casewise Robust Covariance in High Dimensions," arXiv:2505.19925v1 [state.ME].

Ciciretti, V., and Pallotta, A., (2023), "Network Risk Parity: Graph Theory-based Portfolio Construction," *Journal of Asset Management*, 25:136–146.

Chakraborti, A., Hrishidev, Sharma, K., and Pharasi, H., (2020), "Phase Separation and Scaling in Correlation Structures of Financial Markets," *Journal of Physics: Complexity*, 2:015002.

Chatterjee, S., (2021), "A New Coefficient of Correlation," *Journal of the American Statistical Association*, Vol 116(536), 2009-2022.

Chatterjee, S., (2024), "A Survey of Some Recent Developments in Measures of Association," *Probability and Stochastic Processes*, Springer Nature, Singapore.

Chmeilowski, P., (2014), "General Covariance, the Spectrum of Riemannium and a Stress Test Calculation Formula," *Journal of Risk*, 16(6), 1-17.

Choi, J., and Shin, D., (2021), "A self-normalization break test for correlation matrix," *Statistical Papers*, 62(5).

Church, Christ (2012). "The asymmetric t-copula with individual degrees of freedom", Oxford, UK: University of Oxford Master Thesis, 2012.

Colombo, D., and Maathuis, M., (2014), "Order-Independent Constraint-Based Causal Structure Learning," *Journal of Machine Learning Research*, 15, 3921-3962.

Cordoba, I., Varando, G., Bielza, C., and Larranaga, P., (2018), "A fast Metropolis-Hastings method for generating random correlation matrices," *IDEAL*, pp. 117-124, part of Lec Notes in Comp Sci., Vol 11314.

Cota, R., (2019), "Shortfalls of theHierarchical Risk Parity," https://www.linkedin.com/pulse/shortfalls-hierarchical-risk-parity-rafael-nicolas-fermin-cota/

Czado, C., (2025), "Vine Copula Based Structural Equation Models," *Computational Statistics and Data Analysis*, pp.453-477.

Czado, C., and Nagler, T., (2022), "Vine Copula Based Modeling," *Annual Review of Statistics and Its Application*, pp.453-477.

Dalio, R., (2017), Principles of Life and Work, 1st ed., Simon & Schuster.

Dalitz, C., Arning, J., and Goebbels, S., (2024), "A Simple Bias Reduction for Chatterjee's Correlation," arXiv:2312.15496v2.

De Lara, L., (2023), "On the (in)compatibility between potential outcomes and structural causal models and its signification in counterfactual inference," arXiv:2309.05997v3 [math.ST].

Deidda, C., Engelke, S., and De Michele, C., (2023), "Asymmetric Dependence in Hydrological Extremes," *Water Resources Research*, Vol. 59, Issue 12.

Digitale, J., Martin, J., and Glymour, M., (2023), "Tutorial on Directed Acyclic Graphs," *Journal of Clinical Epidemiology*, 142, 264-267.

XI. References

du Plessis, H., and van Rensburg, P., (2020), "Risk-based Portfolio Sensitivity to Covariance Estimation," *Investment Analysts Journal*, 49(3), 243-268.
Ebadi, M., Chenouri, S., Lin, D., and Steiner, S., (2021), "Statistical monitoring of the covariance matrix in multivariate processes: A literature review," *Journal of Quality Technology*, 54(3), 269-289.
Embrechts, P., Hofert, M., and Wang, R., (2016), "Bernoulli and Tail-Dependence Compatibility," *The Annals of Applied Probability*, Vol. 26(3), 1636-1658.
Engle, R., Ledoit, O., and Wolf, M., (2019), "Large dynamic covariance matrices," *Journal of Business & Economic Statistics*, 37(2):363–375.
Epozdemir, M., (2021), "Reverse Stress Testing: A critical assessment tool for risk managers and regulators," S&P Global, BLOG — Aug 10, 2021, https://www.spglobal.com/market-intelligence/en/news-insights/research/reverse-stress-testing-assessment-tool-risk-managers-regulators
European Banking Authority, (2013), REGULATION (EU) No 575/2013 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL, Capital Requirements Regulation (CRR), Articles 375(1), 376(3)(b), 377. https://www.eba.europa.eu/regulation-and-policy/single-rulebook/interactive-single-rulebook/12674
Fang, Q., Jiang, Q., and Qiao, X., (2024), "Large-Scale Multiple Testing of Cross-Covariance Functions with Applications to Functional Network," arXiv:2407.19399v1 [math.ST] 28 Jul.
Felippe, H., Viol, A., de Araujo, D. B., da Luz, M. G. E., Palhano-Fontes, F., Onias, H., Raposo, E. P., and Viswanathan, G. M., (2021), "The von Neumann entropy for the Pearson correlation matrix: A test of the entropic brain hypothesis," working paper, arXiv:2106.05379v1
Felippe, H., Viol, A., de Araujo, D. B., da Luz, M. G. E., Palhano-Fontes, F., Onias, H., Raposo, E. P., and Viswanathan, G. M., (2023), "Threshold-free estimation of entropy from a Pearson matrix," working paper, arXiv:2106.05379v2.
Feng, C., and Zeng, X., (2022), "The Portfolio Diversification Effect of Catastrophe Bonds and the Impact of COVID-19," working paper, https://ssrn.com/abstract=4215258
Fernandez-Duran, J.J., and Gregorio-Dominguez, M.M., (2023), "Testing the Regular Variation Model for Multivariate Extremes with Flexible Circular and Spherical Distributions," arXiv:2309.04948v2.
Fisher, R. A., (1915), "Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population," *Biometrika*, 10, 507-521.
Fisher, R. A., (1921), "On the 'probable error' of a coefficient of correlation deduced from a small sample," *Metron*, 1(4), 1-32.
Fisher, R. A., (1928), "The General Sampling Distribution of the Multiple Correlation Coefficient," Rothamsted Experimental Station, Harpenden, Herts.
Franca, W., and Menegatto, V., (2022), "Positive definite functions on products of metric spaces by integral transforms," *Journal of Mathematical Analysis and Applications*, 514(1).
Galeeva, R., Hoogland, J., & Eydeland, A., (2007), "Measuring Correlation Risk," publicly available manuscript.
Garcin, M., and Nicolas, M.L.D., (2024), "Nonparametric estimator of the tail dependence coefficient: balancing bias and variance," *Statistical Papers*, 65, 4875–4913.
Gao, M., Li, Q., (2024), "A Family of Chatterjee's Correlation Coefficients and Their Properties," arXiv:2403.17670v1 [stat.ME]
Gebelein, H. (1941), "Das statistische problem der korrelation als variations-und eigenwertproblem und sein zusammenhang mit der ausgleichsrechnung," *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift fur Angewandte Mathematik und Mechanik*, 21 (6), 364–379.
Genest, C., Neslehova, J., and Ghorbal, N., (2010), "Spearman's footrule and Gini's gamma: a review with complements," *Journal of Nonparametric Statistics*, 22(8), 937-954.
Ghosh, R., Mallick, B., and Pourahmadi, M., (2021) "Bayesian Estimation of Correlation Matrices of Longitudinal Data," *Bayesian Analysis*, 16, Number 3, pp. 1039–1058.
Gini, C., (1914), "L'Ammontare e la Composizione della Ricchezza delle Nazione," Torino: Bocca.
Golts, M., and Jones, G., "A Sharper Angle on Optimization," ssrn.com, 1483412.
Gong, H., Lu, C., and Zang, Y., (2024), "Distribution-consistency Structural Causal Models" arXiv:2401.15911v2 [cs.AI]
Greenspan, A., (1999), "New Challenges for Monetary Policy," Remarks before a symposium sponsored by the Federal Reserve Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming, August 27, 1999.
Greiner, R. (1909), 'Über das fehlersystem kollektivmaßlehre', *Zeitschrift für Mathematik und Physik* 57, 121–158,225–260,337–373.
Gretton, A., Fukumizu, K., Teo, C., Song, L., Schölkopf, B., and Smola, A. (2007), 'A Kernel Statistical Test of Independence', *Advances in Neural Information Processing Systems*, 20.
Grothe, O., Schnieders, J., and Segers, J., (2014), "Measuring Associatoin and Depenence Between Random Vectors," Journal of Multivariate Analysis, Vol. 123, 96-110.
Gupta, A.K. and Nagar, D.K., (2000), Matrix Variate Distribution, Hall/CRC, Boca Raton.
Hamed, K., (2011), "The distribution of Kendall's tau for testing the significance of cross-correlation in persistent data," *Hydrological Sciences Journal*, 56:5, 841-853.
Han, F., (2021), "On extensions of rank correlation coefficients to multivariate spaces," *Bernoulli News*, 28(2): 7-11.
Hardin, J., Garcia, S., and Golan, D., (2013), "A method for generating realistic correlation matrices," *Annals of Applied Statistics*, 7(3): 1733-1762.
Heller, R., Heller, Y., and Gorfine, M., (2013), 'A Consistent Multivariate Test of Association Based on Ranks of Distances', *Biometrika*, 100(2), 503–510.
Hansen, P., and Luo, Y., (2024), "Robust Estimation of Realized Correlation: New Insight about Intraday Fluctuations in Market Betas," arXiv:2310.19992v1.
Heinen, A., and Valdesogo, A., (2022), "The Kendall and Spearman rank correlations of the bivariate skew normal distribution," *Scandinavian Journal of Statistics*,49:1669–1698.
Heiny, J., and Yao, J., (2022), "Limiting Distributions fro Eigenvalues of Sample Correlation Matrices from Heavy-tailed Populations," arXiv:2003.03857v2 [math.PR].
Hellton, K., (2020), "Penalized Angular Regression for Personalized Predictions," arXiv:2001.09834v2 [stat.ME].
Hernán, M., Hsu, J., and Healy, B., (2019), "A Second Chance to Get Causal Inference Right: A Classification of Data Science Tasks," *Chance*, 32(1), 42-49.
Higham, N., (1988), "Computing the nearest symmetric positive semi-definite matrix," *Linear algebra and its applications*, 103 (1988), 103–118.
Higham, N., (2002), "Computing the nearest correlation matrix – a problem from finance," *IMA Journal of Numerical Analysis*, 22, 329-343.
Hirschfeld, H., (1935), "A connection between correlation and contingency," *Mathematical Proceedings of the Cambridge Philosophical Society*, 31 (4), 520–524.
Hisakado, M. and Kaneko, T., (2023), "Deformation of Marchenko-Pastur distribution for the correlated time series," arXiv:2305.12632v1.
Hochberg, Y., (1988), "A Sharper Bonferroni Procedure for Multiple Tests of Significance, *Biometrika*, 75, 800-802.
Hoeffding, W. (1948). "A Non-parametric Test of Independence." *Annals of Mathematical Statistics*, 19:546–557.
Holzmann, H., and Klar, B., (2024) "Lancaster Correlation - A New Dependence Measure Linked to Maximum Correlation," arXiv:2303.17872v2 [stat.ME].
Hommel, G., (1988), "A Stagewise Rejective Multiple test Procedure Based on a Modified Bonferroni Test," *Biometrika*, 75, 383-386.
Hyvärinen, A., Zhang, K., Shimizu, S., and Hoyer, P., (2010), "Estimation of a Structural Vector Autoregression Model Using Non-Gaussianity," *Journal of Machine Learning Research*, 11, 1709−1731.
Joarder, A., and Ali, M., (1992), "Distribution of the Correlation Matrix for a Class of Elliptical Models," Communications in Statistics – Theory and Methods, 21(7), 1953-1964.
Johnstone, I., (2001), "On the distribution of the largest eigenvalue in principal components analysis," *The Annals of Statistics*, 29(2): 295–327, 2001.
Jondeau, E., (2016), "Asymmetry in Tail Dependence of Equity Portfolios," *Computational Statistics & Data Analysis*, Vol 100, pp351-368.
Jordan, Camille, (1875), "Essai sur la géométrie à dimensions," *Bulletin de la Société Mathématique de France*, 3: 103–174.
Junker. R., Griessenberger, F., and Trutschnig, W., (2021), "Estimating scale-invariant directed dependence of bivariate distributions," *Computational Statistics & Data Analysis*, Volume 153.
Kahn, A., (1962), "Topological Sorting of Large Networks", *Communications of the ACM*, 5 (11), 558–562.
Kelly, B., Malamud, S., Pourmohammadi, M., and Trojani, F., (2024), "Universal Portfolio Shrinkage," NBER Working Paper Series, Working Paper 32004, http://www.nber.org/papers/w32004
Ke, C., (2019), "A New Independence Measure and its Applications in High Dimensional Data Analysis," Doctoral Dissertation, University of Kentucky.
Kendall, M. (1938), "A New Measure of Rank Correlation," *Biometrika*, 30 (1–2), 81–89.
Kim, J., and Finger, C., (1998), "A Stress Test to Incorporate Correlation Breakdown," *Journal of Risk*, 2(3), 5-19.
Kim, W., and Lee, Y., (2016), "A Uniformly Distributed Random Portfolio," *Quantitative Finance*, Vol. 16, No. 2, pp.297-307.

XI. References

Kim, M., Daniels, M., Rooney, W., Willcocks, R., Walter, G., and Vandenborne, K., (2025), "A new algorithm for sampling parameters in a structured correlation matrix with application to estimating optimal combinations of muscles to quantify progression in Duchenne muscular dystrophy," https://arxiv.org/abs/2506.21719v1[stat.ME].

Kiruluta, A., Lemos, A., and Burity, P., (2025), "Operator-Based Machine Learning Intelligence: A Hilbert Space Framework for Spectral Learning and Symbolic Reasoning," arXiv:2507.21189v1 [cs.LG].

Knyazev, A. and Argentati, M., (2002), "Principal Angles between Subspaces in an A-Based Scalar Product: Algorithms and Perturbation Estimates", *SIAM Journal on Scientific Computing*, 23 (6): 2009–2041.

Koike, T., Lin, L., and Wang, R., (2024), "Invariant Correlation Under Marginal Transforms," *Journal of Multivariate Analysis*, 204, 105361.

Krupskii, P., and Joe, H., (2014), "Tail-weighted Measures of Dependence," *Journal of Applied Statistics*, 42(3), 614-629.

Kubiak, S., Weyde, T., Galkin, O., Philps, D., and Gopal, R., (2024), "Denoising Diffusion Probabilistic Model for Realistic Financial Correlation Matrices," *ICAIF '24: Proceedings of the 5th ACM International Conference on AI in Finance*, pp. 1-9. Also see https://github.com/szymkubiak/DDPM-for-Correlation-Matrices

Kurowicka, D., (2014). "Joint Density of Correlations in the Correlation Matrix with Chordal Sparsity Patterns," *Journal of Multivariate Analysis*, 129 (C): 160–170.

Lam, T., Dornemann, N., and Dette, H., (2025), "A New Two-Sample Test for Covariance Matrices in High Dimensions: U-Statistics Meet Leading Eigenvalues," arXiv:2506.06550v1 [math.ST].

Lan, S., Holbrook, A., Elias, G., Fortin, N., Ombao, H., andShahbaba, B. (2020), "Flexible Bayesian Dynamic Modeling of Correlation and Covariance Matrices," *Bayesian Analysis*, 15(4), 1199–1228.

Latif, S., and Morettin, P., (2014), "Estimation of a Spearman-Type Multivariate Measure of Local Dependence," *International Journal of Statistics and Probability*, 3(2).

Lauria, D., Rachev, S., and Trindade, A., (2021), "Global and Tail Dependence: A Differential Geometry Approach," arXiv:2106.05865v1 [stat.AP].

Ledoit, O., and Wolf, M., (2017), "Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks," *The Review of Financial Studies*, 30(12):4349–4388.

Ledoit, O., and Wolf, M., (2022a), "Markowitz portfolios under transaction costs," Working paper series, Department of Economics, (420).

Ledoit, O., and Wolf, M., (2022b), "Quadratic shrinkage for large covariance matrices," *Bernoulli*, 28(3):1519–1547.

Lee, Y., Kim, J., Kim, W., and Fabozzi, F., (2024), "An Overview of Machine Learning for Portfolio Optimization," The Journal of Portfolio Management, 51(2), 131-148.

Lemyre. F., and Quessy, J., (2024), "Kendall's-tau-based Inference for Gradually Changing Dependence Structures," *Statistical Papers*, Volume 65, 2033–2075.

Lewandowski, D.; Kurowicka, D.; Joe, H. (2009). "Generating random correlation matrices based on vines and extended onion method". *Journal of Multivariate Analysis*, 100 (9): 1989–2001.

Leyder, S., Raymaekers, J., and Rousseeuw, P., (2025), "Robust Distance Covariance," *International Statistical Review*, https://doi.org/10.1111/insr.70005.

Li, D., Cerezetti, F., and Cheruvelil, R., (2024), "Correlation breakdowns, spread positions and central counterparty margin models," Journal of Financial Market Infrastructures,11(3), 2049-5404.

Li, Q., (2018), "Covariance Modelling with Hypersphere Decomposition Method and Modified Hypersphere Decomposition Method," Doctoral Thesis, University of Manchester.

Li, G., Zhang, A., Zhang, Q., Wu, D., and Zhan, C., (2022), "Pearson Correlation Coefficient-Based Performance Enhancement of Broad Learning System for Stock Price Prediction," *IEEE Transactions on Circuits and Systems—II: Express Briefs*, Vol 69(5), 2413-2417.

Li, J., Chu, B., Scheller, I., Gagneur, J., and Maathuis, M., (2025), "Root Cause Discovery via Permutations and Cholesky Decomposition," arXiv:2410.12151v5 [stat.ME].

Li, W. ,Yao, J., (2018), "On structure testing for component covariance matrices of a high-dimensional mixture," *Journal of the Royal Statistical Society Series B (Statistical Methodology)*, 80(2):293-318.

Li, X., and Joe, H., (2024), "Multivariate Directional Tail-weighted Dependence Measures," *Journal of Multivariate Analysis*, Vol 203.

Li, Y., (2025), "Large Sample Correlation Matrices with Unbounded Spectrum," *Journal of Multivariate Analysis*, 205, 105373.

Lin, Z., and Han, F., (2023), "On Boosting the Power of Chatterjee's Rank Correlation," *Biometrika*,110(2), 283–299.

Lindskog, F., McNeil, A., Schmock, U., (2003), "Kendall's Tau for Elliptical Distributions," In: Bol, G., Nakhaeizadeh, G., Rachev, S.T., Ridder, T., Vollmer, KH. (eds) Credit Risk. Contributions to Economics, Physica-Verlag HD.

Liu, Y., and Shang, P., (2025), "Differential Distance Correlation and Its Applications," arXiv:2507.00524 [stat.ME].

Loretan, M., and English, W., (2000), "Evaluating Correlation Breakdowns during Periods of Market Volatility," International Finance Discussion Papers (IFDP), https://www.federalreserve.gov/econres/ifdp/evaluating-correlation-breakdowns-during-periods-of-market-volatility.htm and https://www.bis.org/publ/confer08k.pdf

Lu, F., Xue, L., and Wang, Z., (2019), "Triangular Angles Parameterization for the Correlation Matrix of Bivariate Longitudinal Data," *J. of the Korean Statistical Society*, 49:364-388.

Ma, T., Yang, F., Ke, H., and Ren, Z., (2025), "Robust Distance correlation for Variable Screening," *Stat*, 14(3).

MacKinnon, D., and Lamp, S., (2021), "A Unification of Mediator, Confounder, and Collider Effects," *Prev Sci.*, 22(8), 1185-1193.

Madar, V., (2015), "Direct Formulation to Cholesky Decomposition of a General Nonsingular Correlation Matrix," *Statistics & Probability Letters*, Vol 103, pp.142-147.

Makalic, E., Schmidt, D., (2018), "An efficient algorithm for sampling from sin(x)^k for generating random correlation matrices," arXiv: 1809.05212v2 [stat.CO].

Maltsev, A., and Malysheva, S. (2024), "Eigenvalue Statistics of Elliptic Volatility Model with Power-law Tailed Volatility," arXiv:2402.02133v1 [math.PR].

Manistre, J., (2008), "A Practical Concept of Tail Correlation," *Society of Actuaries*.

Marchenko, A., Pastur, L., (1967), "Distribution of eigenvalues for some sets of random matrices," *Matematicheskii Sbornik*, N.S. 72 (114:4): 507–536.

Markowitz, H., (1952), "Portfolio Selection," *The Journal of Finance*, 7, 77-91.

Marti, G., (2019), "CorrGAN: Sampling Realistic Financial Correlation Matrices Using Generative Adversarial Networks," arXiv:1910.09504v1 [q-fin.ST].

Marti, G., Goubet, V., and Nielsen, F., (2021), "cCorrGAN: Conditional Correlation GAN for Learning Empirical Conditional Distributions in the Elliptope," arXiv:2107.10606v1 [q-fin.ST]

Marti, G. (2020) TF 2.0 DCGAN for 100×100 financial correlation matrices [Online]. Available at: https://marti.ai/ml/2019/10/13/tf-dcgan-financial-correlation-matrices.html. (Accessed: 17 Aug 2020)

Martin, C. and Mahoney, M., (2018), "Implicit Self-Regularization in Deep Neural Networks: Evidence from Random Matrix Theory and Implications for Learning," *Journal of Machine Learning Research*, 22 (2021) 1-73.

McNeil, A , Frey, R., and Embrechts, P., (2015), Quantitative risk management: Concepts, techniques and tools, Princeton University Press.

Meckes, M., (2013), "Positive Definite Metric Spaces," arXiv:1012.5863v5 [math.MG].

Meijer, R., Krebs, T., and Goeman, J., (2018), "Hommel's Procedure in Learn Time," *Biometrical Journal*, 61(1), 73-82.

Metsämuuronen, J., (2022), "Reminder of the Directional Nature of the Product-Moment Correlation Coefficient," *Academia Letters*, Article 5313.

Meucci, A., (2010a), "The Black-Litterman Approach: Original Model and Extensions," The Encyclopedia of Quantitative Finance, Wiley, 2010

Meucci, A., (2010b), "Fully Flexible Views: Theory and Practice," arXiv:1012.2848v1

Mitchell, C., Ryne, R., and Hwang, K., (2022), "Using kernel-based statistical distance to study the dynamics of charged particle beams in particle-based simulation codes," *Phys. Rev. E*, 106, 065302.

Mograby, G., (2025), "Hierarchical Minimum Variance Portfolios: A Theoretical and Algorithmic Approach," arXiv:2503.12328v1 [q-fin.PM].

Muirhead, R., (1982), Aspects of Multivariate Statistical Theory, Wiley Interscience, Hoboken, New Jersey.

Nawroth, A., Anfuso, F. and Akesson, F., (2014), "Correlation Breakdown and the Influence of Correlations on VaR," https://ssrn.com/abstract=2425515 or http://dx.doi.org/10.2139/ssrn.2425515

Ng, F., Li, W., and Yu, P., (2014), "A Black-Litterman Approach to Correlation Stress Testing," *Quantitative Finance*, 14:9, 1643-1649.

Niu, L., Liu, X., and Zhao, J., (2020), "Robust Estimator of the Correlation Matrix with Sparse Kronecker Structure for a High-Dimensional Matrix-variate," *Journal of Multivariate Analysis*, 177, 104598.

Opdyke, JD, (2003), "Fast Permutation Tests That Maximize Power Under Conventional Monte Carlo Sampling for Pairwise and Multiple Comparisons," *Journal of Modern Applied Statistical Methods*, 2(1), 27-49.

Opdyke, JD, (2011), "Permutation Tests (and Sampling Without Replacement) Orders of Magnitude Faster Using SAS®," *InterStat*, January, 2011.

Opdyke, JD, (2013), "Bootstraps, Permutation Tests, and Sampling Orders of Magnitude Faster Using SAS®," *Computational Statistics: WIREs Interdisciplinary Reviews*, 5(5), 391-405.

Opdyke, JD, (2022), "Beating the Correlation Breakdown: Robust Inference and Flexible Scenarios and Stress Testing for Financial Portfolios," QuantMindsEdge: Alpha and Quant Investing: New Research: Applying Machine Learning Techniques to Alpha Generation Models, June 6 (available at www.DataMineit.com).

XI. References

Opdyke, JD, (2023), "Beating the Correlation Breakdown: Robust Inference and Flexible Scenarios and Stress Testing for Financial Portfolios," Columbia University, NYC–School of Professional Studies: Machine Learning for Risk Management, Invited Guest Lecture, March 20 (available at www.DataMineit.com).

Opdyke, JD, (2024a), Keynote Presentation: "Beating the Correlation Breakdown, for Pearson's and Beyond: Robust Inference and Flexible Scenarios and Stress Testing for Financial Portfolios," QuantStrats11, NYC, March 12, 2024 (available at www.DataMineit.com).

Opdyke, JD, (2024b), RoundTable Writeup: "Association-based vs Causal Research: the Hype, the Contrasts, and the Stronger-than-expected Complementary Overlaps," QuantStrats11, NYC, March 12, 2024 (available at www.DataMineit.com).

Packham, N., and Woebbeking, F., (2023), "Correlation scenarios and correlation stress testing," *Journal of Economic Behavior & Organization*, Vol 205, pp.55-67.

Pafka, S., and Kondor, I., (2004), "Estimated correlation matrices and portfolio optimization," Physica A: Statistical Mechanics and its Applications, Vol 343, 623-634.

Palomar, D., (2025), Portfolio Optimization: Theory and Application, Cambridge University Press.

Papenbrock, J., Schwendner, P., Jaeger, M., and Krugel, S., (2021), "Matrix Evolutions: Synthetic Correlations and Explainable Machine Learning for Constructing Robust Investment Portfolios," *Journal of Financial Data Science*, 51-69.

Parlatore, C. and Philippon, T., (2024), "Designing Stress Scenarios," NBER Working Paper No. w29901.

Pascual-Marqui, R., Kochi, K., and Kinoshita, T., (2024), "Distance-based Chatterjee correlation: a new generalized robust measure of directed association for multivariate real and complex-valued data: arXiv:2406.16458 [stat.ME].

Pearl, J., (2009), "Causal inference in statistics: An overview," *Statistics Surveys*, 3:96–146, 2009.

Pearson, K., (1895), "VII. Note on regression and inheritance in the case of two parents," *Proceedings of the Royal Society of London*, 58: 240–242.

Pfitzinger, J., and Katzke, N., (2025), "Perspectives on Optimal Hierarchical Portfolio Selection," https://ssrn.com/abstract=4409173

Pham-Gia, T., Choulakian, V., (2014), "Distribution of the Sample Correlation Matrix and Applications," *Open Journal of Statistics*, 4(5).

Pinheiro, J., and Bates, D., (1996), "Unconstrained parametrizations for variance-covariance matrices," *Statistics and Computing*, Volume 6, 289–296.

Polakow, D., Gebbie, T., and Flint, E., (2023), "Epistemic Limits of Empirical Finance: Causal Reductionism and Self-Reference," arXiv:2311.16570v2 [q-fin.GN]

Pramanik, P., (2024), "Measuring Asymmetric Tails Under Copula Distributions," *European Journal of Statistics*, 4:7.

Puccetti, G., (2022), "Measuring linear correlation between random vectors," *Information Sciences*, Vol 607, 1328-1347.

Quessy, J., Meriem, S., and Favre, A., (2013), "Multivariate Kendall's Tau for Change-Point Detection in Copulas," *The Canadian Journal of Statistics*, 41(1), 1-25.

Qian, E. and Gorman, S. (2001). "Conditional Distribution in Portfolio Theory." *Financial Analysts Journal*, 44-51.

Qin, T., and Wei-Min, H., (2024), "Epanechnikov Variational Autoencoder," arXiv:2405.12783v1 [stat.ML] 21 May 2024.

Ramsey, J., Glymour, M., Sanchez-Romero, R., and Glymour, C., (2017), "A million variables and more: the fast greedy equivalence search algorithm for learning high-dimensional graphical causal models, with an application to functional magnetic resonance images," *International Journal of Data Science and Analytics*, 3, 121–129.

Rebonato, R., and P. Jäckel, P., (2000) "The Most General Methodology to Create a Valid Correlation Matrix for Risk Management and Option Pricing Purposes," *Journal of Risk*, 2(2), 17-27.

Reddi, S., Ramdas, A., Poczos, B., Singh, A., and Wasserman, L., (2015), "On Decreasing Power of Kernel and Distance based Nonparametric Hypothesis Tests in High Dimensions," arXiv:1406.2083v2 [stat.ML].

Reichenbach, H., (1956). *The Direction of Time*, University of California Press.

Reisach, A., Suárez, A., Weichwald, S., and Chambaz, A., (2025), "The Case for Time in Causal DAGs," arXiv:2501.19311v3 [stat.ME] July 8, 2025.

Rodgers, J., and Nicewander, A., (1988), The American Statistician, 42(1), 59-66.

Rodriguez Dominguez, A., and Yadav, O., (2024), "Measuring causality with the variability of the largest eigenvalue," *Data Science in Finance and Economics*.

Rodriguez Dominguez, A., (2023), "Portfolio optimization based on neural networks sensitivities from assets dynamics respect common drivers," *Machine Learning with Applications*, 11(15), 100447.

Rodriguez Dominguez, A., (2025a), "Causal Portfolio Optimization: Principles and Sensitivity-Based Solutions," arXiv:2504.05743v2 [q-fin.PM].

Rodriguez Dominguez, A., (2025b), "Is Causality Necessary for Efficient Portfolios? A Computational Perspective on Predictive Validity and Model Misspecification," arXiv:2507.23138v3 [q-fin.PM].

Rubsamen, R., (2023), "Random Correlation Matrices Generation," https://github.com/lequant40/random-correlation-matrices-generation

Ruelloux, P., Bongiorno, C., and Challet, D., (2025), "Noise-Proofing Universal Portfolio Shrinkage," arXiv:2511.10478v1 [q-fin.RM]

Sabato, S., Yom-Tov, E., Tsherniak, A., Rosset, S., (2007), "Analyzing systemlogs: A new view of what's important," *Proceedings, 2nd Workshop of Computing Systems ML*, pp.1–7.

Saxena, S., Bhat, C., and Pinjari, A., (2023), "Separation-based parameterization strategies for estimation of restricted covariance matrices in multivariate model systems," *Journal of Choice Modelling*, Vol. 47, 100411.

Schmidt, R. and Stadtmüller, U., (2006), "Non-parametric estimation of tail dependence," *Scandinavian Journal of Statistics*, 33(2):307-335.

Sejdinovic, D., Sriperumbudur, B., Gretton, A., and Fukumizu, K., (2013) "Equivalence of Distance-Based and RKHS-Based Statistics in Hypothesis Testing," *The Annals of Statistics*, 41(5), 2263-2291.

Sheppard, W., (1899), "On the application of the theory of error to cases of normal distribution and normal correlation," *Philosophical Transactions of the Royal Society of London (A)*, 92, 101–167.

Shimizu, S., Inazumi, T., Sogawa, Y., Hyvarinen, A., Kawahara, Y., Washio, T., Hoyer, P., and Bollen, K., (2011), "DirectLiNGAM: A Direct Method for Learning a Linear Non-Gaussian Structural Equation Model," *Journal of Machine Learning Research*, 12,1225-1248.

Shmueli, G., (2010), "To Explain or to Predict?," *Statistical Science*, 25(3), 289-310.

Siburg, K., Strothmann, C., and Weiß, G., (2024), "Comparing and quantifying tail dependence," *Insurance: Mathematics and Economics*, Vol 118, 95-103.

Spearman, C., (1904), "'General Intelligence,' Objectively Determined and Measured," *The American Journal of Psychology*, 15(2), 201–292.

Stollenwerk, M., (2025), "Probability Distributions for Realized Covariance Measures," *Journal of Econometrics*, https://doi.org/10.1016/j.econom.2025.105954.

Sun, F., and Huang, X., (2025), "Application of Regularized Covariance Matrices in Logistic Regression and Portfolio Optimization," Scientific Reports, 15(1).

Szekely, G., Rizzo, M., and Bakirov, N., (2007), "Measuring and Testing Dependence by Correlation of Distances," *The Annals of Statistics*, 35(6), pp2769-2794.

Taleb, N., (2007), The Black Swan: The Impact of the Highly Improbable, Random House, Inc, New York, New York.

Taraldsen, G. (2021), "The Confidence Density for Correlation," *The Indian Journal of Statistics*, 2021.

Thakkar, A., Patel, D., and Shah, P., (2021), "Pearson Correlation Coefficient-based performance enhancement of Vanilla Neural Network for Stock Trend Prediction," *Neural Computing and Applications*, 33:16985-17000.

Tripathi, Y., Chatla, S., Chang, Y., Huang, L., and Shiefh, G., (2022), "A Nonlinear Correlation Measure with Applications to Gene Expression Data," *PLoS ONE*, 17(6): e0270270.

Trucíos Maza, Carlos César, (2025), "Hierarchical risk clustering versus traditional risk-based portfolios: an empirical out-of-sample comparison," https://ssrn.com/abstract=5247627 or http://dx.doi.org/10.2139/ssrn.5247627

Tsay, R., and Pourahmadi, M., (2017), "Modelling structured correlation matrices," *Biometrika*, 104(1), 237-242.

Tumminello, M., Aset, T., Di Matteo, T., and Mantegna, R., (2005), "A tool for filtering information in complex systems," *Proceedings of the National Academy of Sciences*, 102(30), 10421-10426.

van den Heuvel, E., and Zhan, Z., (2022), "Myths About Linear and Monotonic Associations: Pearson's r, Spearman's ρ, and Kendall's τ," *The American Statistician*, 76:1, 44-52.

Vanni, F., Hitaj, A., and Mastrogiacomo, E., (2024), "Enhancing Portfolio Allocation: A Random Matrix Theory Perspective," *Mathematics*, 12(9), 1389

Veleva, E., (2017), "Generation of Correlation Matrices," *AIP Conference Proceedings 1895*, 1230008, https://doi.org/10.1063/1.5007425

Vinod, H., (2022), "Asymmetric Dependence Measurement and Testing," arXiv:2211.16645v1 [stat.ME].

Vorobets, A., (2024), "Sequential Entropy Pooling Heuristics," https://ssrn.com/abstract=3936392 or http://dx.doi.org/10.2139/ssrn.3936392

Vorobets, A., (2025), "Portfolio Construction and Risk Management," https://ssrn.com/abstract=4807200 or http://dx.doi.org/10.2139/ssrn.4807200

Wahba, G., (2017), "Emanuel Parzen and a Tale of Two Kernels," *Technical Report No. 1183*, University of Wisconsin, Madison.

Wang, Z, Wu, Y., and Chu, H., (2018), "On equivalence of the LKJ distribution and the restricted Wishart distribution," arXiv:1809.04746v1.

XI. References

Wang, B., Xu, F., and Shu, L., (2019), "A Bayesian Approach to Diagnosing Covariance Matrix Shifts," *Quality and Reliability Engineering International*, 36(2).
Wang, J., Zhu, T., and Zhang, J., (2025), "Test of the Equality of Several High-Dimensional Covariance Matrices: A Normal-Reference Approach," *MDPI: Mathematics*, 13, 295.
Welsch, R., and Zhou, X., (2007), "Application of Robust Statistics to Asset Allocation Models," *REVSTAT– Statistical Journal*, Volume 5(1), 97–114.
Westfall, P., and Young, S., (1993), Resampling Based Multiple Testing, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, New York.
Wright, S., (1921), "Correlation and Causation," *Journal of Agricultural Research*, 20(7), 557-585.
Wright, S., (1992), "Adjusted *P*-Values for Simultaneous Inference," *Biometrics*, 48, 1005-1013.
Xia, L., Cao, R., Du, J., and Chen, X., (2024), "The Improved Correlation Coefficient of Chatterjee," *Journal of Nonparametric Statistics*, pp1-17.
Yu, P., Li, W., Ng, F., (2014), "Formulating Hypothetical Scenarios in Correlation Stress Testing via a Bayesian Framework," *The North Amer. J. of Econ. and Finance*, Vol 27, 17-33.
Yu, S., Alesiani, F., Yu, X., Jenssen, R., and Principe, J., (2021), "Measuring Dependence with Matrix-based Entropy Functional," *The Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI-21)*.
Yu, C., Zhu, Z., and Zhu, K., (2025), "Tensor Dynamic Conditional Correlation Model: A New Way To Pursuit 'Holy Grail of Investing'," arXiv:2502.13461v1 [q-fin.PM]
Yule, G.U., (1897), "On the Significance of Bravais Formulæ for Regression, in the case of skew correlation," Proceedings of The Royal Society London, 477-489.
Zanga, A., Ozkirimli, E., and Stella, F., (2025), "A Survey on Causal Discovery: Theory and Practice," arXiv:2305.10032v2 [cs.AI]
Zar, J., (1999), Biostatistical Analysis, 4th Ed., Prentice Hall, New Jersey.
Zhang, Q., (2024a), "On relationships between Chatterjee's and Spearman's correlation coefficients," Communications in Statistics – Theory and Methods, 54(1), 259-279.
Zhang, Q., (2024b), "On the Extensions of the Chatterjee-Spearman Test" arXiv:2406.16859v1 [stat.ME]
Zhang, Y., (2022), "Angle Based Dependence Measures in Metric Space," arXiv:2206.01459v1 [stat.ME].
Zhang, Y., and Yang, S., (2023), "Kernel Angle Dependence Measures for Complex Objects," arXiv:2206.01459v2
Zhang, W., Leng, C., and Tang, Y., (2015), "A Joint Modeling Approach for Longitudinal Studies," *Journal of the Royal Stat. Society, Series B*, 77(1), 219-238.
Zhang, Y., Tao, J., Yin, Z., and Wang, G., (2022), "Improved Large Covariance Matrix Estimation Based on Efficient Convex Combination and Its Application in Portfolio Optimization," *Mathematics*, 10(22), 4282.
Zhang, Y., Tao, J., Lv, Y., and Wang, G., (2023), "An Improved DCC Model Based on Large-Dimensional Covariance Matrices Estimation and Its Applications," *Symmetry*, 15, 953.
Zhang, G., Jiang, D., and Yao, F., (2024), "Covariance Test and Universal Bootstrap by Operator Norm," arXiv:2412.20019v1 [math.ST].
Zhangshuang, S., Gao, X., Luo, K., Bai, Y., Tao, J., and Wang, G., (2025), "Enhancing High-Dimensional Dynamic Conditional Angular Correlation Model Based on GARCH Family Models: Comparative Performance Analysis for Portfolio Optimization," *Finance Research Letters*, 106808.
Zhao, T., Roeder, K., and Liu, H., (2014), "Positive Semidefinite Rank-based Correlation Matrix Estimation with Application to Semiparametric Graph Estimation," *Journal of Computational Graphics and Statistics*, 23(4):895–922.
Zheng, S., Shi, N.-Z., and Zhang, Z. (2012), "Generalized Measures of Correlation for Asymmetry, Nonlinearity, and Beyond," *Journal of the American Statistical Association*, 107, 1239–1252.
Zheng, X., Aragam, B., Ravikumar, P., and Xing, E., (2018), "Dags with no tears: Continuous optimization for structure learning," *Advances in Neural Information Processing Systems*, 9472–9483.