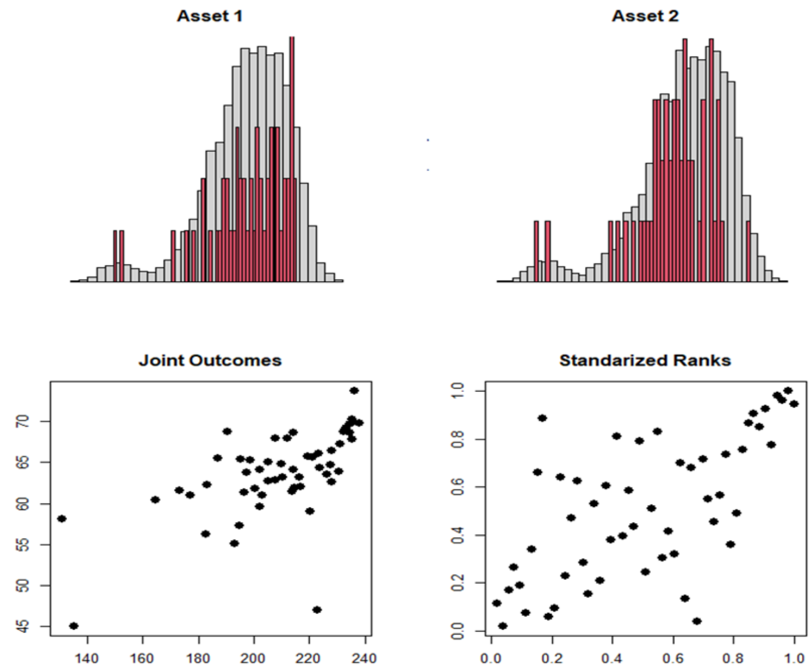
SIMULATING MULTIVARIATE VARIABLES USING NON-PARMAMETRIC SMOOTHED HISTORICAL DATA PROCEDURES

Joseph Atwood

Montana State University - Bozeman



Overview

- **We wish to assess the risk of a static or optimized portfolio of M assets.**

- **Let $\boldsymbol{Y} = [\boldsymbol{y_1}, \boldsymbol{y_2}, \dots, \boldsymbol{y_M}]$ be a sample of n joint historical observations on M variables with $\boldsymbol{Y \sim MV(MS, DS)}$ with marginal structure MS and dependency structure DS.**

- **Common (McNeil, Frey, Embrechts (2025), Hull (2018)) for financial institutions to use the data in the Y matrix with various "Historical" approaches to estimate potential future VaR and cVaR levels.**

Historical Procedures

- **Advantages.**
  - **Nonparametric and distribution free.**
  - **Does not require the estimation/modeling of each asset's marginal distribution.**
  - **Does not require the estimation/specification of the dependency structure.**
  - **Relatively easy to use.**
  - **Can use subsamples to estimate stressed VaR and cVaR.**
- **Disadvantages.**
  - **Future outcomes may differ from past outcomes.**
  - **Both marginals and dependencies tend to vary over time**
  - **See other discussions by Butler and Schachter (1996)**

Historical Procedures

- **We present a "non-parametric" hybrid "historical-modeling" approach that utilizes historical data Y but generates marginal outcomes and dependencies not observed in the historical data.**
    - **Non-parametric estimation of and simulating from each marginal using kernel procedures similar to Butler and Schachter (1996)**
    - **Binds the marginal's together using "historical copula"**
    - **Historical copulas are related to empirical copulas but are (in our opinion) more flexible and often generate more consistent results.**

Marginal Simulation Procedures

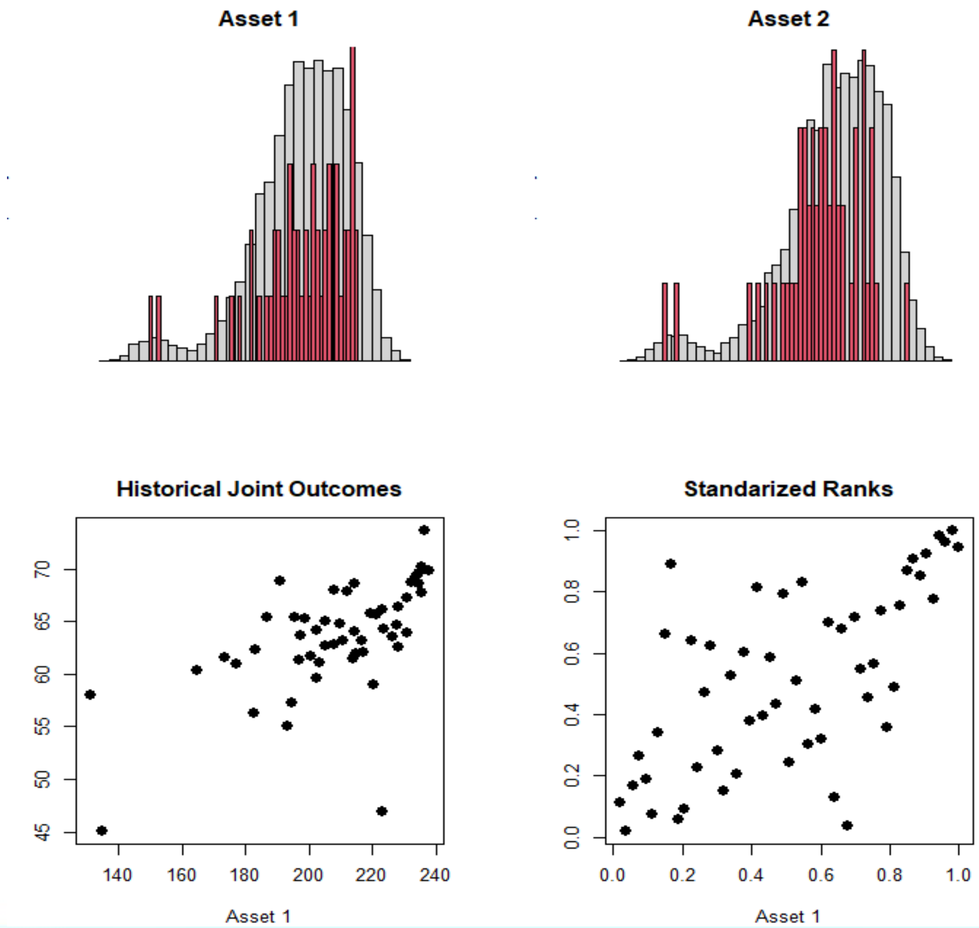
- **Due to time constraints we are going to "flash through" the procedures with some R scripts and examples.**
  - **We want N >= 10000 joint potential outcomes for each of M variables**
  - **Historical Y has n = 53 observations**
  - **nreps = floor(N/n) + 1 = 189 for this example**
  - **N = nreps * n = 10017 for this example**

<u>R script</u>

```
YI = matrix(0,N,M) # Simulated Data Matrix
 for(j in 1:M){
   tmp = density(Y[,j])
   YI[, j] = sample(tmp$x, N, prob=tmp$y,replace=T)
} # end loop on j
```

Data & Simulated Marginal

(Note the Empirical Dependency Structure)



Construct the “Historical Copula”

$$HCOP = \begin{bmatrix} Y \\ Y \\ \vdots \\ Y \end{bmatrix}$$

- Note: HCOP is not strictly a copula but we will use variations of HCOP to reorder the ranks in the simulated YI data.

- If N is a multiple of n, Y and HCOP will have the same marginal and cross-sectional dependency structure and “correlations”.

- Each “block” of HCOP will also have the same marginal and across variable serial dependency as Y

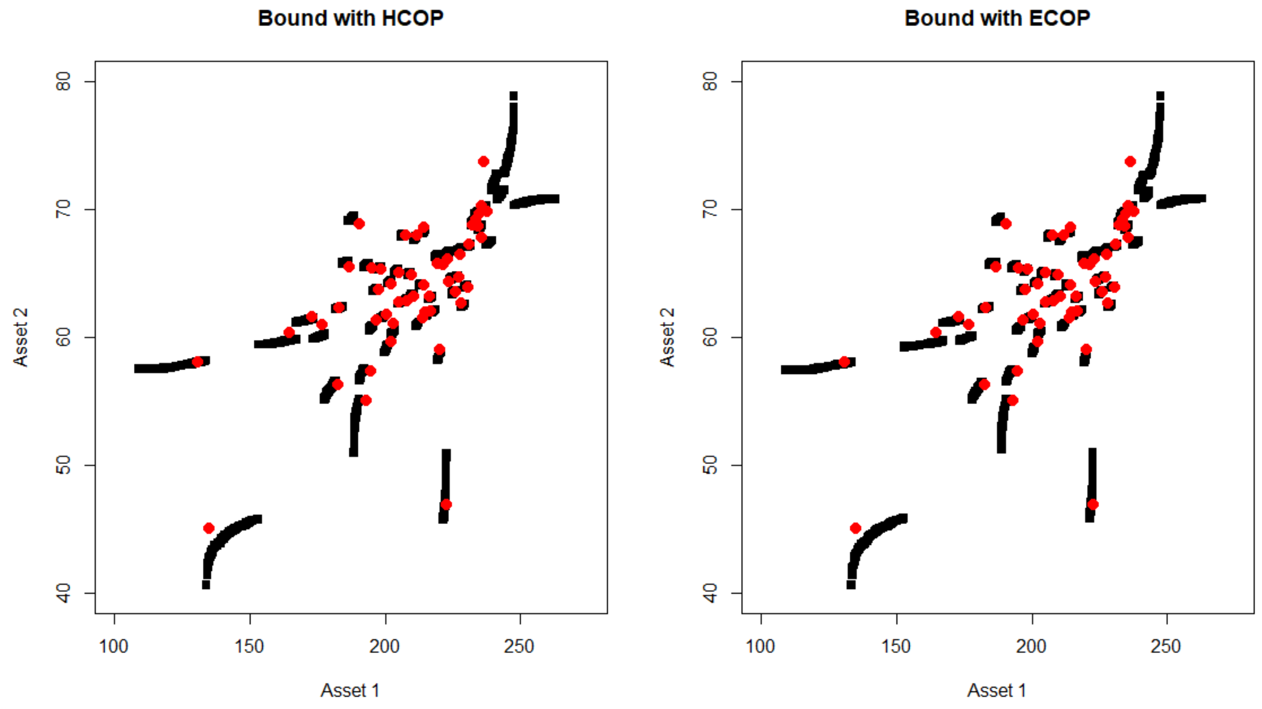
Using HCOP

```
rankorder = function(Y, RM){
  …  # some error and dimension checking
  for(j in 1:ncol(Y){
    ys = sort(Y[,j])
    Y[,j] = ys[rank(RM[,j], ties.method='first')]
  }   # end loop on j
 Y
} # end function rankorder
```

YD = rankorder(YI,HCOP)

Using HCOP

Without smoothing HCOP, using HCOP to bind the YI marginals is equivalent to using the empirical copula in the R package ‘copula’ (Hofert, Kojadadinovic, Maechler, Yan) (2025)



Smoothing HCOP

HCOP can be smoothed is numerous ways.

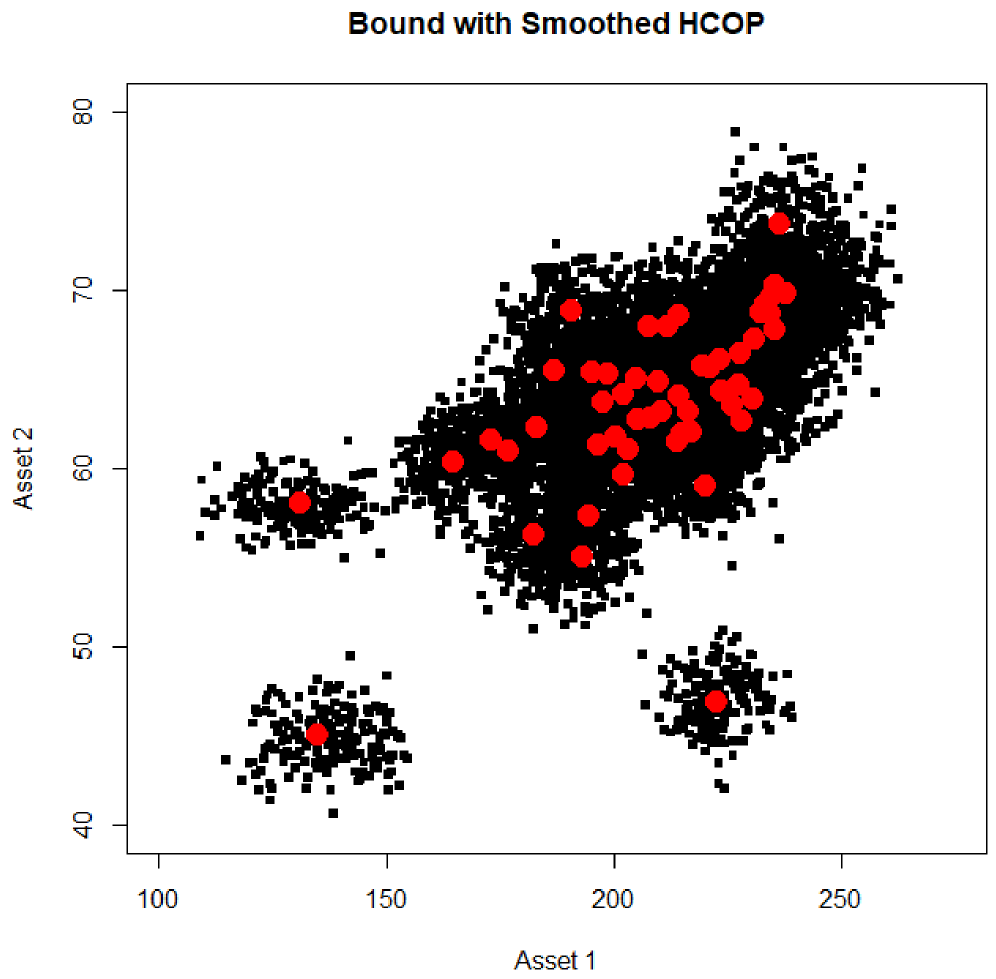
The following uses independent normal variates for smoothing

R script:

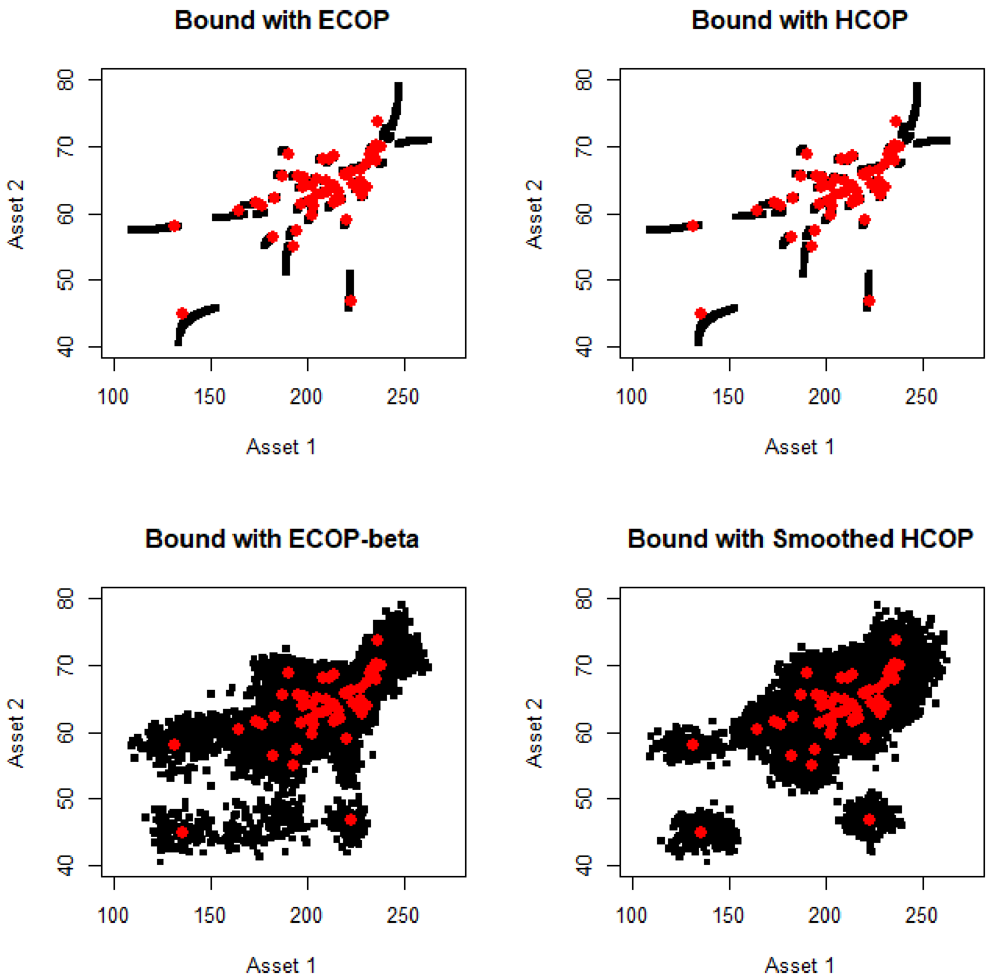
```
sdevs = apply(HCOP,2,sd)
sscale = 0.25
SHCOP = HCOP
for(j in 1:ncol(SHCOP)) {
  SHCOP[,j] = HCOP[,j] + sscale*sdevs[j]*rorm(N)
}  # end loop on j

YDS = rankorder(YI,SHCOP)
```

Binding with Smoothed HCOP



Smoothed HCOP vs Smoothed ECOP



HCOP and Multivariate Normality

HCOP can be used with large number M of marginals to implement ImanConover type process (i.e. bind YI with a normal copula) without having to estimate and decompose the covariance matrix or multiply matrices)

R script:

```
HCOPN = HCOP; J = 100
for(j in 2:J){
 HCOPN = HCOPN = HCOP[sample(1:N,N,replace=F),]
} # end loop on J
HCOPN = HCOPN/sqrt(J)  # not strictly needed
YDN = rankorder(YI,HCOPN)
```

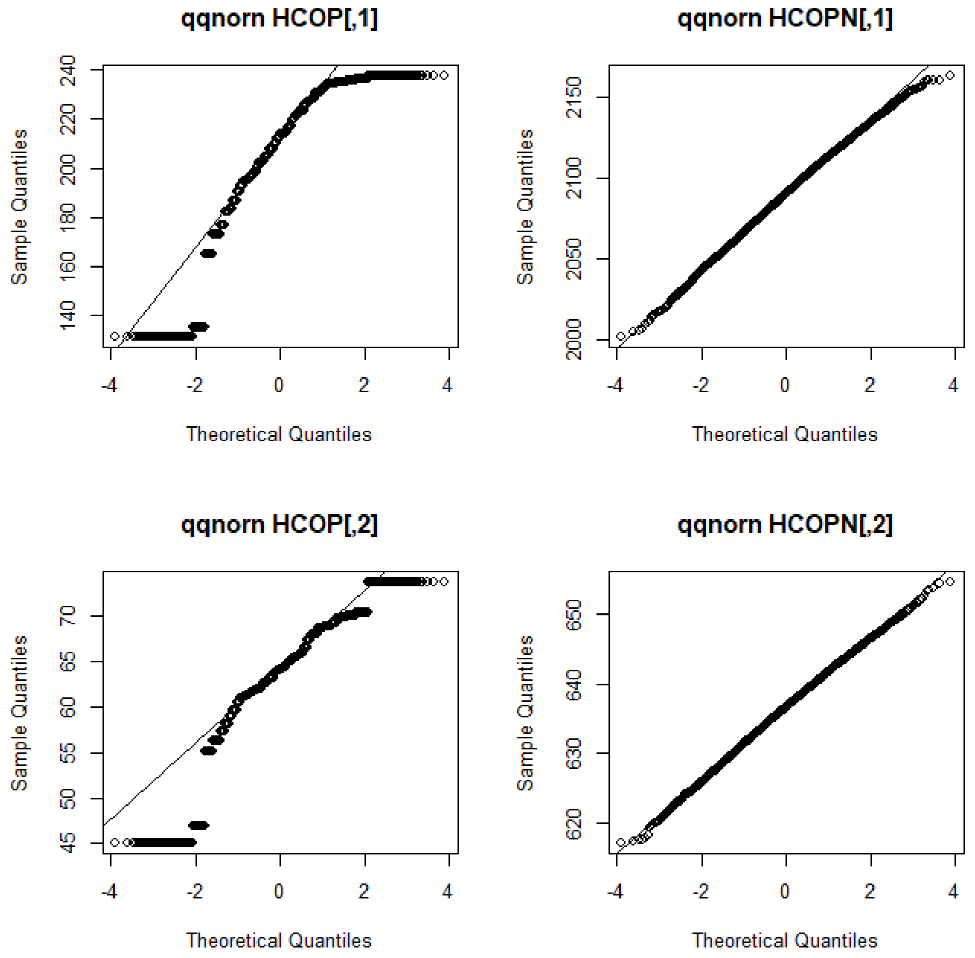
HCOP and Multivariate Normality

By the Multivariate Central Limit Theorem

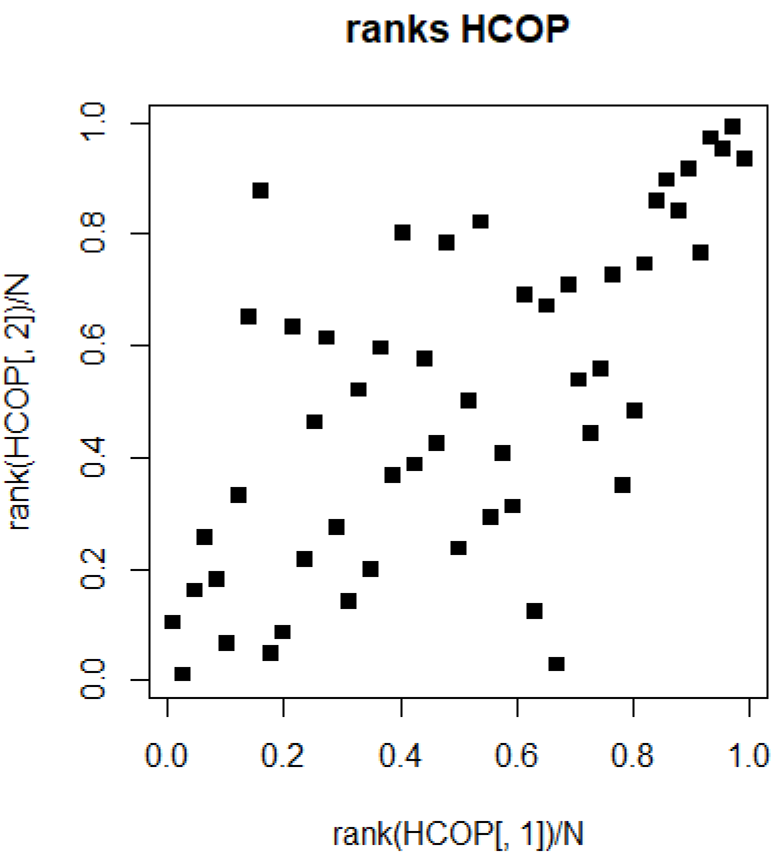
HCOPN ~ MVN(mean(Y), cov(Y)) since Y and HCOP have the same mean and covariance matrix.

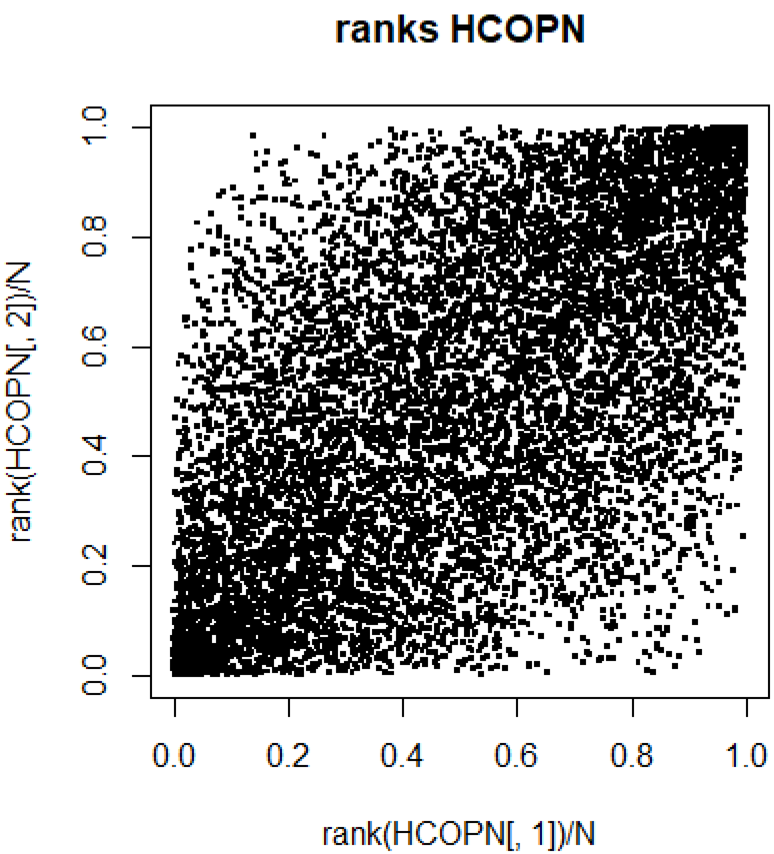
As a result the marginals in YDN will be bound together with a normal copula with correlations cor(Y).

HCOP and HCOPN qqnorm plots



HCOP and HCOPN Rank Plots





HCOP and HCOPN M Asset Times

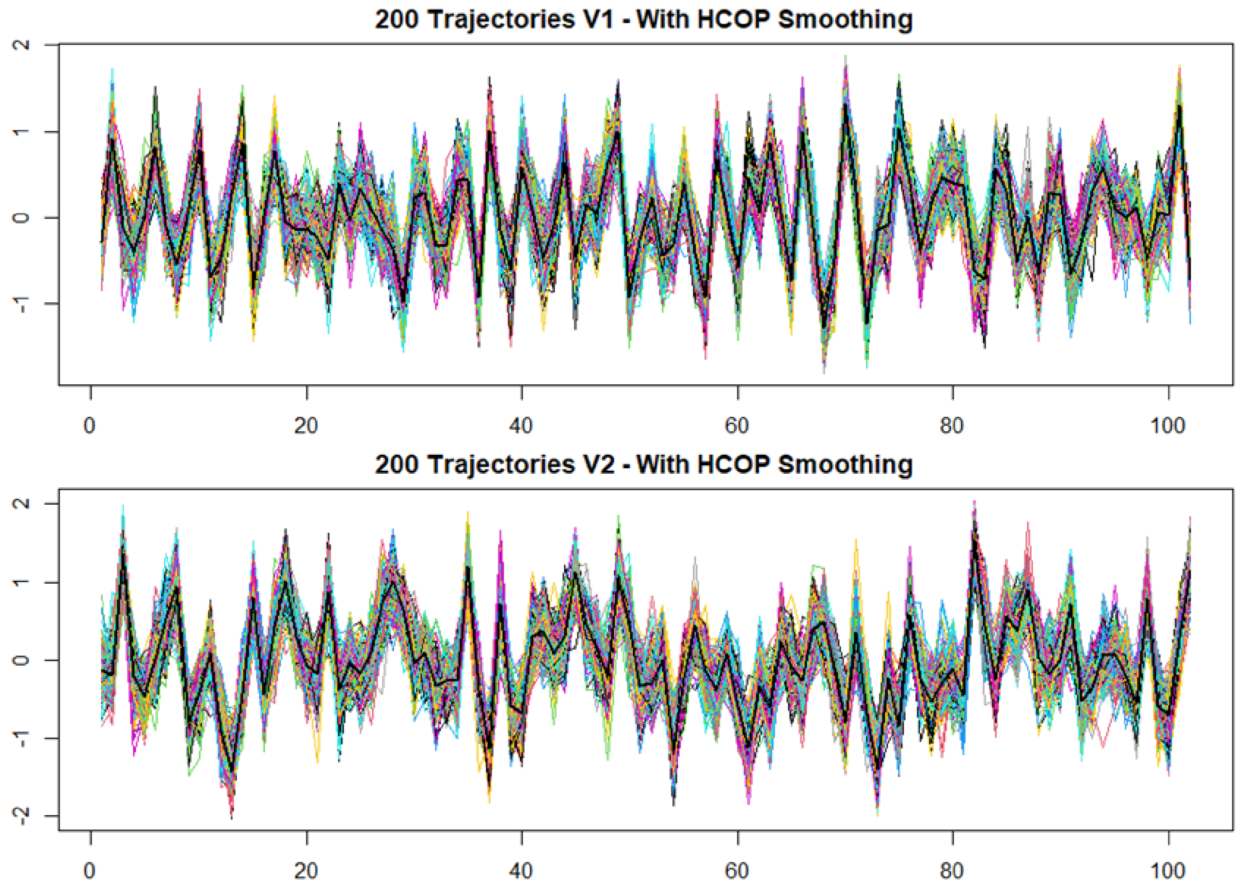
HCOP Computation Times to Estimate and Simulate 10000 Joint Observations of M Variables Using Marginal Kernal Smoothing and HCOP Procedures.

M	1000	2000	4000	8000	16000	32000
Seconds	2.34	4.56	10.27	17.95	36.5	73.98

Computation Times to Simulate 10000 Correlated Normal Observations of M Variables Using Spectral / Cholesky Decomposition Versus Modified HCOP Procedures. (Using open.blas with R)

	Computation Times in Seconds					
M	1000	2000	4000	8000	16000	32000
Spectral/Choletsky	0.7	1.6	6.7	45.4	370.1	2893.6
HCOP	1.4	3.1	6.1	11.6	24.8	50.7

Smoothed HCOP and Serial Dependency



Smoothed HCOP and Cross Sectional/Serial Dependency

HCOP procedures can also be utilized in various "structural applications" to non- parametrically model and simulate potential joint errors from multi equation, panel, time series, or vector autoregressive regression models.

Using HCOP

HCOP procedures are:

- **Largely non-parametric but can be combined with parametric models.**
- **Can accommodate large number M of joint variables.**
- **Fast: Time grows linearly with M**
- **Can be used with MV Normal Copula binding while preserving but not having to compute all** $(M^2 - M)/2$ **historical correlations.**
- **Flexible**
- **Monograph and Code available from author**

References

Atwood, J. (2025) ***Modeling Multivariate Variables using Smoothed Historical Data and Non-Parametric Procedures.*** **Working paper. Department of Agricultural Economics and Economics. Montana State University. Bozeman, MT**

Butler, J. and B. Schachter. (1996) ***Improving Value-at-Risk Estimates by Combining Kernal Estimation with Historical Simulation*****. Office of the Comptroller of the Currency, E&PA Working Paper 96-1, August 1996**

Hull, J. (2023). ***Risk Management and Financial Institutions (6th)*****. John Wiley & Sons. Hoboken, NJ.**

McNeil. A.J., Frey, R. and Embrechts, P. (2015) ***Quantitative Risk Management Concepts, Techniques and Tools (Revised)*****. Princeton University Press, Princeton, NJ.**