Supplementary material to *An information-theoretic approach to build hypergraphs in psychometrics*

Daniele Marinazzo[1], Jan Van Roozendaal[1], Fernando E. Rosas[2,3,4,5], Massimo Stella[6], Renzo Comolatti[7], Nigel Colenbier[1,8], Sebastiano Stramaglia[9,10], and Yves Rosseel[1]

[1]Department of Data Analysis, Ghent University, Belgium

[2]Data Science Institute, Imperial College London, London, UK

[3]Centre for Psychedelic Research, Imperial College London, London, UK

[4]Centre for Complexity Science, Imperial College London, London, UK

[5]Department of Informatics, University of Sussex, Brighton, UK

[6]CogNosco Lab, Dipartimento di Psicologia e Scienze Cognitive, Università di Trento, Rovereto, Italy

[7]Department of Biomedical and Clinical Sciences "L. Sacco", Università degli Studi di Milano, Italy

[8]IRCCS San Camillo Hospital, Venice, Italy

[9]Physics Department, Universitá degli Studi di Bari Aldo Moro, Bari, Italy

[10]INFN Sezione di Bari, Italy

Supplementary material to *An information-theoretic approach to build hypergraphs in psychometrics*

Appendix A - Reanalysis of the Golino *redundancy* framework

Here we analyze the datasets simulated in Christensen et al. (2020), Golino, Christensen, et al. (2021), and Golino, Moulder, et al. (2021) with 2000 observations, three latent factors (factor loadings = 0.4 and three variables per latent factor, without ($c = 0$, 1) and with ($c = 0.3$, 2) correlation between factors. Only redundancy emerges from the analysis with our framework.

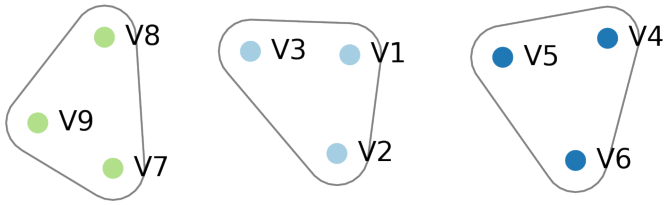

Figure 1

Significant redundant multiplets when the factors are not correlated.

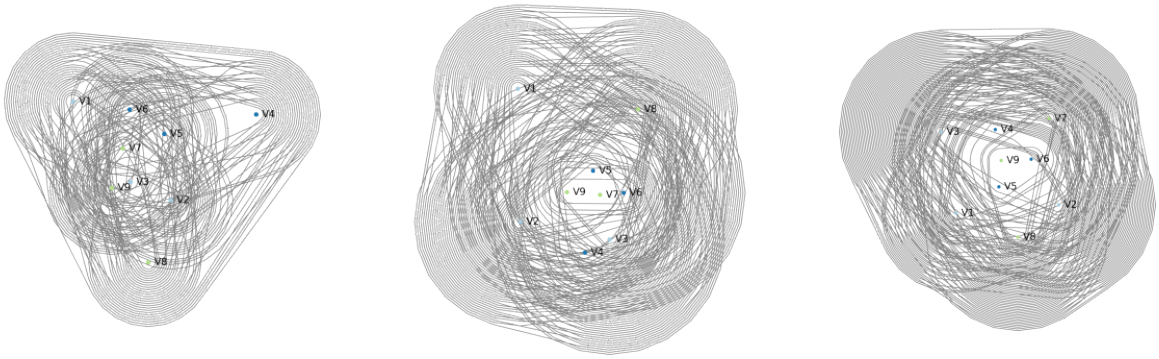

Figure 2

Significant redundant multiplets when the factors are correlated.

Appendix B - A simple example of multiplicative interaction, resulting in a synergistic triplet

.

The code below (in R syntax) shows how one can generate a set of three variables according to a "higher order mechanism" such as a multiplicative interaction, which is classified as a synergistic triplet by the approach presented here. Crucially, this synergistic triplet would be indistinguishable from a triplet that one would obtain from Gaussian variables with a given covariance structure (like those generated according to model 1) and reported in the bottom right panel of figure **??**.

```
# generate x1 and x2, with correlation of 0.2
N <- 2000
Sigma <- matrix(c(1, 0.2, 0.2, 1), 2, 2)
X <- MASS::mvrnorm(n = N, Sigma = Sigma, mu = rep(0, 2))
# extract x1 and x2
x1 <- X[,1]
x2 <- X[,2]
# create triplet with NO interaction
y <- 0.2*x1 + 0.4*x2 + rnorm(N, sd = 1)
Data.no.int <- data.frame(y = y, x1 = x1, x2 = x2)
# correlation for the triplet y/x1/x2:
cor(Data.no.int)
# create two-way interaction term
x1.x2 <- x1 * x2
# create y as a function of x1, x2 AND x1.x2
y <- 0.5*x1 + 0.7*x2 + 0.4*x1.x2
# quadruplet Q contains all information: y, x1, x2 AND x1.x2
Q <- data.frame(y = y, x1 = x1, x2 = x2, x1.x2 = x1.x2)
# triplet T that is fed to the O-information algorithm
T<-Data[,c("y","x1","x2")]
```

References

Christensen, A. P., Garrido, L. E., & Golino, H. F. (2020). Unique variable analysis: A novel approach for detecting redundant variables in multivariate data.

Golino, H. F., Christensen, A. P., Moulder, R., Kim, S., & Boker, S. M. (2021). Modeling latent topics in social media using dynamic exploratory graph analysis: The case of the right-wing and left-wing trolls in the 2016 us elections. Psychometrika, 1–32.

Golino, H. F., Moulder, R., Shi, D., Christensen, A. P., Garrido, L. E., Nieto, M. D., Nesselroade, J., Sadana, R., Thiyagarajan, J. A., & Boker, S. M. (2021). Entropy fit indices: New fit measures for assessing the structure and dimensionality of multiple latent variables. Multivariate Behavioral Research, 56(6), 874–902.