

Discussion of “Graphical models for extremes” by Sebastian Engelke and Adrien S. Hitz

Raphaël Huser¹ and Daniela Cisneros¹

February 26, 2020

We congratulate the authors on this impressive and timely paper, which makes a giant step forward in the modeling and inference for high-dimensional extremes, opening avenues for future research in causal inference and machine learning for extremes. State-of-the-art methods in statistics of extremes were so far limited to very low dimensions when modeling maxima (Castruccio *et al.*, 2016; Huser *et al.*, 2019) and moderately high dimensions when modeling high threshold exceedances (de Fondeville and Davison, 2018). By defining notions such as “*conditional independence for multivariate Pareto distributions (MPDs)*” and “*extremal graphical models*”, Engelke and Hitz have developed a new framework for building and learning sparse models for multivariate extremes in potentially very high dimensions.

Our comments mostly revolve around the extension of the proposed methodology to the spatial extremes context. A crucial assumption in Theorem 1 is graph decomposability. While this assumption includes important classes (e.g., trees and block graphs), it is fairly restrictive for spatial data. For example, when spatial locations form a regular lattice in \mathbb{R}^2 , any non-trivial tree structure violates spatial stationarity. The most natural spatially-structured Markov random field on a regular lattice would assign the same distribution to all first-order neighbors and assume conditional independence for all other pairs of variables (i.e., the edge set only contains first-order neighbors). Such a graph is however *not* decomposable, and it is currently unclear to us if/how the results can be extended to this case. However, to retain computational tractability, a promising approach might be to consider finite mixtures of tree-based MPDs similar to Yu *et al.* (2017) and Vettori *et al.* (2019).

¹King Abdullah University of Science and Technology (KAUST), Computer, Electrical and Mathematical Sciences and Engineering (CEMSE) Division, Thuwal 23955-6900, Saudi Arabia.
E-mails: raphael.huser@kaust.edu.sa; daniela.cisnerosarce@kaust.edu.sa

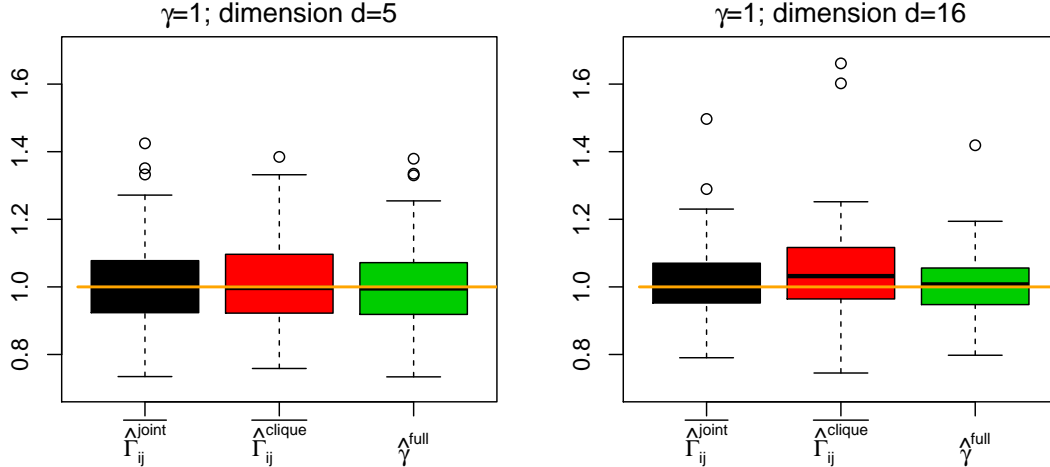


Figure 1: Boxplots of estimates of γ based on the three different estimators mentioned in the text, for trees in dimension $d = 5$ (left) and $d = 16$ (right). The true value $\gamma = 1$ is represented by orange horizontal lines.

A way to impose some spatial structure for data on a regular lattice modeled with tree-based MPDs is to assume the *same* dependence parameter γ for each clique. In this context, the efficiency of the proposed clique-wise inference approach is no longer obvious. To assess this, we replicated the simulation study in §5.5, now specifying the same value $\gamma = \Gamma_{12} = \Gamma_{13} = \Gamma_{24} = \Gamma_{25}$ for each clique. For illustration, we chose $\gamma = 1$. We then estimated γ with the average $\widehat{\Gamma}_{ij} = (\widehat{\Gamma}_{12} + \widehat{\Gamma}_{13} + \widehat{\Gamma}_{24} + \widehat{\Gamma}_{25})/4$, obtained by maximizing either the joint or the clique-wise censored likelihoods (treating the Γ_{ij} s as distinct free parameters), and by fitting the “full” model enforcing the same value γ on all cliques $\{i, j\}$. The average of clique-wise estimates, $\widehat{\Gamma}_{ij}^{\text{clique}}$, is comparable to using a pairwise likelihood with pairs chosen as the cliques themselves. We also explored a larger 16-dimensional tree structure defined on a 4-by-4 spatial lattice. The results, reported in Figure 1, suggest that the estimator $\widehat{\Gamma}_{ij}^{\text{clique}}$ performs almost as well as $\widehat{\gamma}^{\text{full}}$ (based on the “full” model), though the former has a larger loss of efficiency compared to the latter in higher dimensions (about 9% when $d = 5$ versus 33% when $d = 16$, in terms of root mean squared error). Nevertheless, the proposed approach is much faster and still quite accurate in this context.

References

- Castruccio, S., Huser, R. and Genton, M. G. (2016) High-order composite likelihood inference for max-stable distributions and processes. *Journal of Computational and Graphical Statistics* **25**(4), 1212–1229.
- de Fondeville, R. and Davison, A. C. (2018) High-dimensional peaks-over-threshold inference. *Biometrika* **105**(3), 575–592.
- Huser, R., Dombry, C., Ribatet, M. and Genton, M. G. (2019) Full likelihood inference for max-stable data. *Stat* **8**(1), e218.
- Vettori, S., Huser, R., Segers, J. and Genton, M. G. (2019) Bayesian model averaging over tree-based dependence structures for multivariate extremes. *Journal of Computational and Graphical Statistics* To appear.
- Yu, H., Uy, W. and Dauwels, J. (2017) Modeling spatial extremes via ensemble-of-trees of pairwise copulas. *IEEE Transactions on Signal Processing* **65**, 571–586.