

# Conditional coefficient for extremal graphical models

CAPEL Alexandre

2024-12-18

## Contents

<b>1</b>	<b>Extremal/Asymptotical independence</b>	<b>1</b>
1.1	Case $d = 2$ . . . . .	1
1.2	Case $d > 2$ . . . . .	2
<b>2</b>	<b>The conditional extremal coefficient <math>\chi_{ij}^A</math></b>	<b>2</b>
2.1	The trivariate coefficient $\chi_{ijk}$ . . . . .	3
2.2	Case where $A = \{k\}$ . . . . .	3
2.3	Definition for any $A$ . . . . .	3
2.4	Simulation study for the Hüsler-Reiss model . . . . .	4
<b>3</b>	<b>Proofs</b>	<b>9</b>
3.1	Alternative writing for the trivariate coefficient . . . . .	9
3.2	Inequalities of section 2 . . . . .	10

Let  $V = \{1, \dots, d\}$  the set of nodes of a graph  $\mathcal{G} = (V, E)$ .

We want to define a new coefficient used for characterize the conditional dependence to apply for a pairwise Markov property in the extremal graphical model setting (Engelke and Hitz 2020).

## 1 Extremal/Asymptotical independence

Let  $X$  a random  $d$ -dimensional vector where the marginal distribution are standard Frechet.

### 1.1 Case $d = 2$

One famous quantity in extreme value theory is the bivariate extremal coefficient which is defined as follow :

$$\chi_{1,2} = \lim_{q \rightarrow 1} \mathbb{P}(F(X_1) > q \mid F(X_2) > q) = \lim_{q \rightarrow 1} \chi_{1,2}(q)$$

if the right-hand term exists, where  $F$  is the distribution function of a standard Frechet.

We say that  $X_1$  and  $X_2$  are asymptotically independent if  $\chi_{1,2} = 0$  and asymptotically dependent otherwise.

If  $X$  is in the attraction domain of generalized Pareto distribution  $Y$  of exponent measure  $\Lambda$ , there is an equivalence between :

$$\chi_{1,2} = 0 \iff \left( \forall x = (x_1, x_2) > 0, H(x) = H_1(x_1) + H_2(x_2) \right)$$

where  $H(y) = \Lambda[(\mathbb{R}_+)^2 \setminus [0, y]]$ .

And we know in (Strokorb 2020) that is equivalent to the extended extremal independence  $Y_1 \perp\!\!\!\perp_e Y_2$ .

## 1.2 Case $d > 2$

For some dimension  $d$ , we have similar results. Let  $A$  and  $B$  a partition of  $V$ . With the extended extremal independence of (Strokorb 2020), one can show a similar equivalence as above where :

$$Y_A \perp\!\!\!\perp_e Y_B \iff \left( \forall x = (x_A, x_B) > 0, H(x) = H_A(x_A) + H_B(x_B) \right)$$

We can also define the Sum of Extremal COefficient of the partition  $\{A, B\}$  by :

$$SECO_{A,B} = H_A(\mathbb{1}_A) + H_B(\mathbb{1}_B) - H(\mathbb{1})$$

where  $H_C$  is the exponent measure function of the extreme limit of the sub-vector  $X_C$ .

With (Ferreira 2011), we can show that  $SECO_{A,B} = 0$  if and only if we have  $Y_A \perp\!\!\!\perp_e Y_B$ .

### Motivation

For both situation, we can stress that there is a list of equivalent assertion which are following the same structure : extremal (or asymptomatic) independence, exponent measure decomposition, and coefficient equality.

So we can wonder, in the context of extremal conditional independence, if we can define a conditional exponent measure function  $H^C$  to have the relation :

$$Y_A \perp\!\!\!\perp_e Y_B \mid Y_C \iff \left( \forall x = (x_A, x_B) > 0, H^C(x) = H_A^C(x_A) + H_B^C(x_B) \right)$$

where  $A, B, C$  is a partition of  $V$ .

Furthermore, can we find some coefficient  $\theta_{AB}^C$  (calculable from data) and a constant  $c \in \mathbb{R}$  such that :

$$\theta_{AB}^C = c \iff Y_A \perp\!\!\!\perp_e Y_B \mid Y_C$$

*More I look at these relations, more I think it will be difficult to get... The previous relation of independence where strongly linked to the independence of the corresponding max stable distribution. But in our case, we can't talk about conditional independence for the max stable distribution because it does not exist (I mean, it is equivalent to the classical independence).*

*If I really want to build these relations, I think we will need to use or create new notions, which ideally can be linked to the previous ones (exponent measure, etc...).*

## 2 The conditional extremal coefficient $\chi_{ij}^A$

The aim of the section is to present a new coefficient for extreme graphical models and try to link the general case to the univariate one, which is simpler to compute. As above, let  $X$  a random  $d$ -dimensional vector where the marginal distribution are standard Frechet.

## 2.1 The trivariate coefficient $\chi_{ijk}$

We define the trivariate extremal coefficient as :

$$\chi_{ijk} = \lim_{q \rightarrow 1} \mathbb{P}(F(X_i) > q, F(X_j) > q \mid F(X_k) > q) = \lim_{q \rightarrow 1} \chi_{ijk}(q)$$

if the right-hand term exists.

In that case, if the vector  $X$  is in the attraction domain of  $Y$ , the corresponding generalized Pareto distribution, then we also get :

$$\chi_{ijk} = \mathbb{P}(Y_i > 1, Y_j > 1 \mid Y_k > 1) \quad (1)$$

We would like to use this coefficient in order to characterize the conditional independence (or dependence) between the component in a similar sense of the extremal conditional independence (see (Engelke and Hitz 2020)).

## 2.2 Case where $A = \{k\}$

By analogy on the extremal coefficient  $\chi_{ij}$ , we define the conditional extremal coefficient by :

$$\chi_{ij}^k = \lim_{q \rightarrow 1} \mathbb{P}(F(X_i) > q \mid F(X_j) > q, F(X_k) > q) = \frac{\chi_{ijk}}{\chi_{jk}}$$

in the same setting as above.

Of course, the equation in the last section gives us another way to compute this quantity using the corresponding generalised Pareto distribution  $Y$  :

$$\chi_{ij}^k = \mathbb{P}(Y_i > 1 \mid Y_j > 1, Y_k > 1)$$

## 2.3 Definition for any $A$

Now, let's define for some  $A \subset V$ .

Let  $i, j \in V$ , then the conditional extremal coefficient is defined as :

$$\chi_{ij}^A = \lim_{q \rightarrow 1} \mathbb{P}(F(X_i) > q \mid F(X_j) > q, \|F(X_A)\|_\infty > q) = \lim_{q \rightarrow 1} \chi_{ij}^A(q)$$

where  $F(X_A) = (F(X_k))_{k \in A}$ , if the right-hand limit exists.

We could use different extension for a general subset  $A$  of  $V$  as for example  $\lim_{q \rightarrow 1} \mathbb{P}(F(X_i) > q \mid F(X_j) > q, F(X_A) > q)$ , but let's focus on the first one before.

First, we can write another expression for this coefficient, similar to the previous section, which is :

$$\chi_{ij}^A = \mathbb{P}(Y_i > 1 \mid Y_j > 1, \|Y_A\|_\infty > 1)$$

(the demonstration is the same as the case where  $A = \{k\}$ )

More, we have some inequalities between the conditional coefficient  $\chi_{ij}^A$  and all sub-coefficient  $\chi_{ij}^k$  for  $k \in A$ .

**Proposition.** Let  $A \subset V$  and  $i, j \notin A$ . Then we have :

- $\chi_{ij}^k \leq |A| \chi_{ij}^A$  for all  $k \in A$ .
- $\chi_{ij}^A \leq \sum_{k \in A} \chi_{ij}^k$ .

In particular, we have thanks to these inequalities, the equivalence :

$$\left( \forall k \in A, \chi_{ij}^k = 0 \right) \iff \chi_{ij}^A = 0$$

An appreciated properties will be that a conditional independence between two nodes occurs if and only if,  $\chi_{ij}^A = 0$  with  $A = V \setminus \{i, j\}$  and so  $\chi_{ij}^k = 0$  for all  $k \neq i, j$ .

## 2.4 Simulation study for the Hüsler-Reiss model

We would like to test the last assumption in the case of a Hüsler-Reiss model, where the conditional independence can be verified easily. For  $d = 3$  for example, let's define the variogram matrix of our Hüsler-Reiss model :

$$\Gamma = \begin{pmatrix} 0 & a & b \\ a & 0 & c \\ b & c & 0 \end{pmatrix}$$

Then, thanks to the extended precision matrix  $\Theta$  defined in (Hentschel, Engelke, and Segers 2023) we have that :

- $Y_1 \perp\!\!\!\perp Y_2 | Y_3 \iff a = b + c$
- $Y_1 \perp\!\!\!\perp Y_3 | Y_2 \iff b = a + c$
- $Y_2 \perp\!\!\!\perp Y_3 | Y_1 \iff c = a + b$

More, in the Hüsler-Reiss model, the extremal coefficient (bivariate and trivariate) can be computed fastly using :

- $\chi_{ij} = 2 - 2\Phi(\sqrt{\Gamma_{ij}}/2)$
- $\chi_{ijk} = 3 - 2\Phi(\sqrt{\Gamma_{ij}}/2) - 2\Phi(\sqrt{\Gamma_{jk}}/2) - 2\Phi(\sqrt{\Gamma_{ik}}/2) + \sum_{k=1}^3 \Phi_2((\sqrt{\Gamma_{ik}}/2)_{i \neq k}; A_k \Sigma^{(k)} A_k)$  with  $A_k = \text{diag}((\sqrt{\Gamma_{ik}}^{-1})_{i \neq k})$  and where  $\Phi$  is the distribution function of a standard Gaussian variable and  $\Phi_2(\cdot; \Sigma)$  the distribution function of a gaussian vector of covariance matrix  $\Sigma$ .

We produce a random matrix restricted by the relation  $a = b + c$  (to get a graph with no node linking 1 and 2). The coefficients  $b$  and  $c$  are simulated with by an uniform distribution on  $[0, 5]$ .

```
tar_load(Gamma_cond_indep_12)      # load our simulated variogram
print(Gamma_cond_indep_12)
```

```
##           [,1]      [,2]      [,3]
## [1,] 5.75331 4.73967 1.01364
```

Of course, we have to check the conditional negative definiteness of the matrix (to know if the matrix is a possible Hüsler-Reiss variogram). In order to do that we use the `graphicalExtremes`.

```
checkGamma(to_matrix(Gamma_cond_indep_12), returnBoolean = T)    # function checking the validity
```

```
## [1] TRUE
```

Now, we are able to compute the trivariate coefficient, and consequently check if this is zero or not.

```
tar_load(chi_trivariate_example)    # load the value of the trivariate coef for our matrix
print(chi_trivariate_example)
```

```
## [1] 0.2001413
```

So finally it is not 0...

This is not very surprising because the condition that we wanted is about a coefficient which is totally symmetric between  $i, j$  and  $k$ , so the conditional independence (in the  $d = 3$  case at least) of two implies the conditional independence for all and then the graph is empty.

But, let's check another thing. We are going to simulate several matrix and see what is the trend for the value of the conditional extremal coefficient.

### Simulation with $n$ replicates

Here, let's check for our example for a tri-dimensional Hüsler-Reiss graphical model with no edge between nodes 1 and 2 if there we can tell something else with different variogram. For our setup :

- simulate  $n = 2000$  matrices
- the coefficient  $b$  and  $c$  are produced by a uniform random variable in the interval  $[0, 25]$ .

For the conditional extremal coefficient, we get these results :

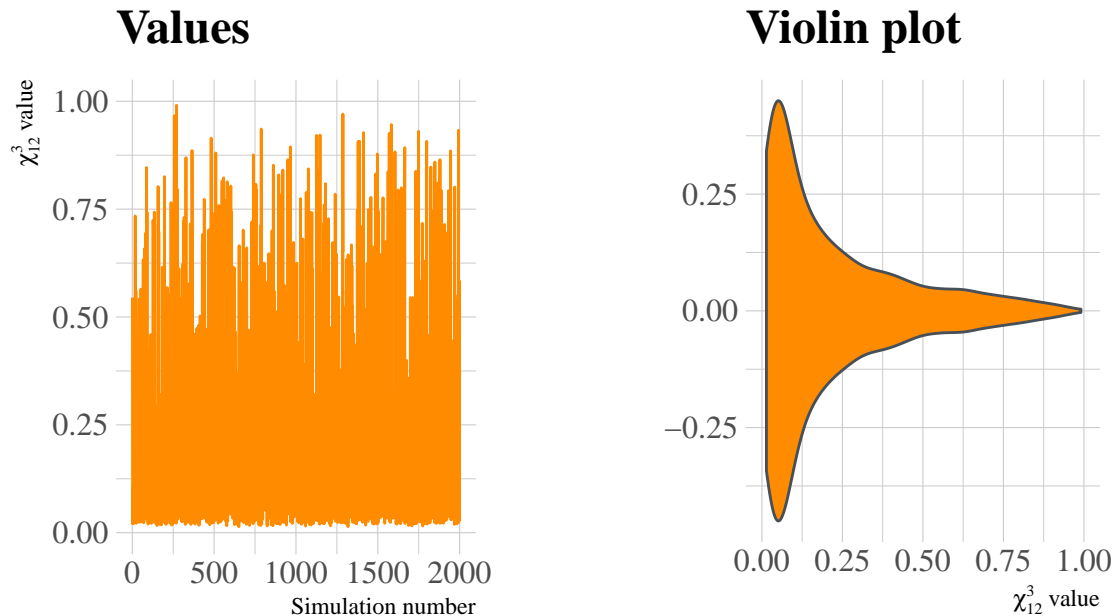


Figure 1: Conditional coefficient computed with the simulations

We can not have any conclusion with such figures... Perhaps, we notice a concentration near 0, 1 but nothing else. If we focus our attention in the trivariate coefficient, we will obtain the following figure :

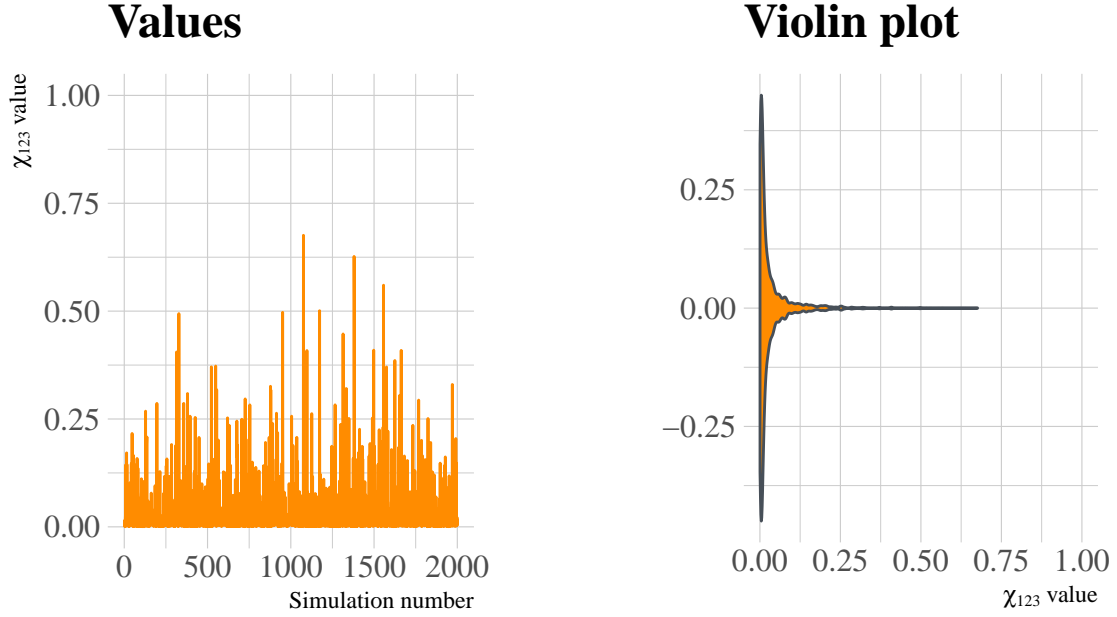


Figure 2: Trivariate coefficient computed with the simulations

Again, we can not say anything, even if the concentration around 0 is a little bit stronger. We can explain this by the fact that the trivariate coefficient is lower than the conditional one.

However, under the conditional independence assumption, one can notice that :

$$\chi_{123} = \mathbb{P}(Y_1 > 1, Y_2 > 1 \mid Y_3 > 1) = \mathbb{P}(Y_1 > 1 \mid Y_3 > 1) \mathbb{P}(Y_2 > 1 \mid Y_3 > 1) = \chi_{1,3} \chi_{2,3}$$

and so that the quotient coefficient  $\frac{\chi_{123}}{\chi_{1,3} \chi_{2,3}} = 1$ .

*Finally, it is not... But there is what I've done which indicates the equality is false. I will explain why below.*

Let's see if it is the case for our previous simulations. We obtained :

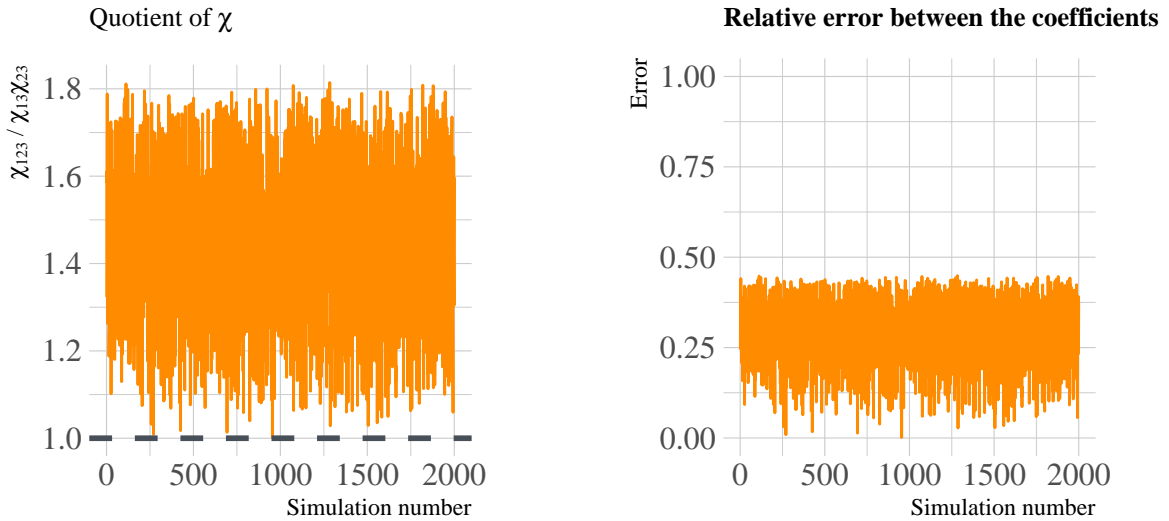


Figure 3: Computation of the quotient coefficient (left, the expected value 1 is the grey dashed line) and gap between the two coefficients (right).

We can notice that is not really what we expected... It seems that there is some error during the computation, but if we try with others values for the coefficient of the variogram that there are close to zero, we notice better more accurate results as it should be :

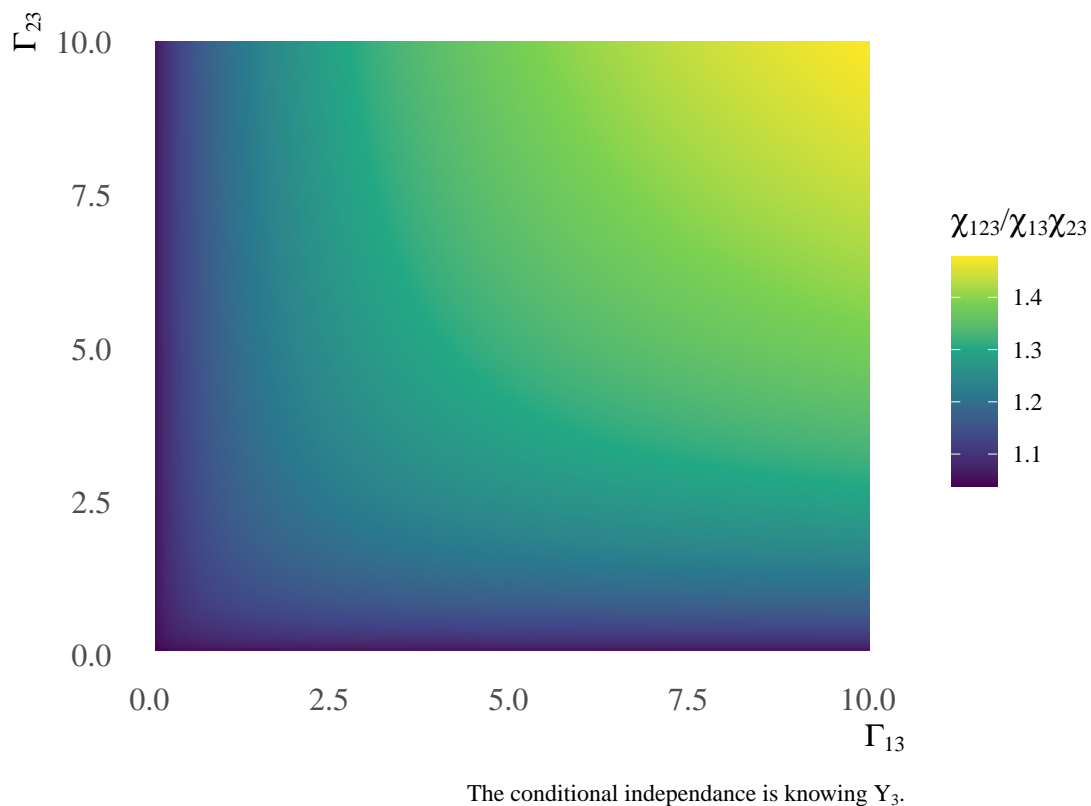
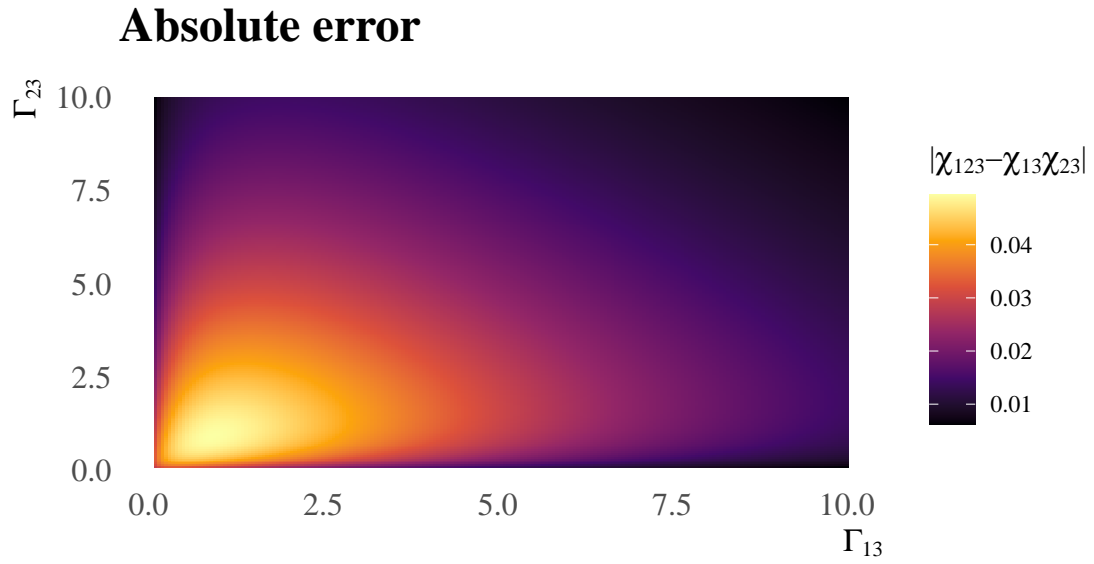
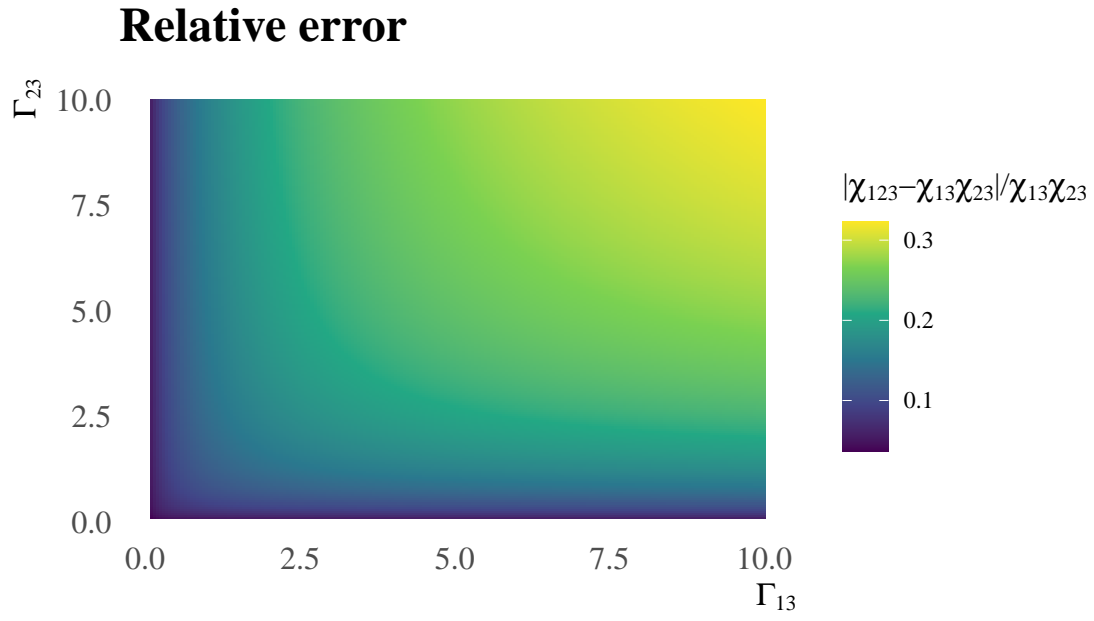


Figure 4: Quotient value for each possible matrix when conditional independence

Perhaps, there is an effect of the fact that our values are very small. To see this effect, let's build the absolute and relative error for the coefficient computation and see what we can say.



The conditional independence is knowing  $Y_3$ .

Figure 5: Different error types for each possible matrix when conditional independence



Actually, we made a mistake. The relation of independence we wrote above was a pure independence relationship, not conditional dependence... With another notation it corresponds to :

$$\mathbb{P}(Y_1^3 > 1, Y_2^3 > 1) = \mathbb{P}(Y_1^3 > 1) \times \mathbb{P}(Y_2^3 > 1)$$

which is true when  $Y_1^3 \perp\!\!\!\perp Y_2^3$ , not when  $Y_1^3 \perp\!\!\!\perp Y_2^3 \mid Y_3^3$ .

*To get the relation, except using the factorization of the density, I don't know really how to formulate the conditional independence using probabilities because the notion of conditional probabilities given a random variable is very weird and not that intuitive (because it's conditional given a particular  $\sigma$ -algebra)...*

## 3 Proofs

### 3.1 Alternative writing for the trivariate coefficient

We will show that we can compute the trivariate coefficient with the formula :

$$3 - \Lambda_{ij}(1, 1) - \Lambda_{ik}(1, 1) - \Lambda_{jk}(1, 1) + \Lambda_{ijk}(1, 1, 1)$$

and the two writing can be expressed with the latter.

First, let's take a look to the limit side :

$$\begin{aligned} \chi_{ijk}(q) &= \frac{1}{1-q} [1 - 3q + \mathbb{P}(F(X_i) < q, F(X_j) < q) + \mathbb{P}(F(X_i) < q, F(X_k) < q) \\ &\quad + \mathbb{P}(F(X_j) < q, F(X_k) < q) - \mathbb{P}(F(X_i) < q, F(X_j) < q, F(X_k) < q)] \\ &= 3 - \frac{1 - \mathbb{P}(F(X_i) < q, F(X_j) < q)}{1-q} - \frac{1 - \mathbb{P}(F(X_i) < q, F(X_k) < q)}{1-q} \\ &\quad - \frac{1 - \mathbb{P}(F(X_j) < q, F(X_k) < q)}{1-q} + \frac{1 - \mathbb{P}(F(X_i) < q, F(X_j) < q, F(X_k) < q)}{1-q} \\ &= 3 - \frac{1}{F^{-1}(q)} \left[ \frac{F^{-1}(q)(1 - \mathbb{P}(X_{ij} < F^{-1}(q)))}{1-q} - \frac{F^{-1}(q)(1 - \mathbb{P}(X_{ik} < F^{-1}(q)))}{1-q} \right. \\ &\quad \left. - \frac{F^{-1}(q)(1 - \mathbb{P}(X_{jk} < F^{-1}(q)))}{1-q} + \frac{F^{-1}(q)(1 - \mathbb{P}(X_{ijk} < F^{-1}(q)))}{1-q} \right] \end{aligned}$$

But, as  $X$  is a random vector with standard Frechet marginals,  $X$  is a regular variation random vector, that means that :

$$\lim_{u \rightarrow \infty} u(1 - \mathbb{P}(X_C < u)) = \Lambda_C(\mathbb{1}_C)$$

where  $\Lambda_C$  is the exponent measure of the marginal extreme limit of  $X_C$ .

But, we have  $F^{-1}(q) \rightarrow \infty$  and  $-\log(q) \sim 1 - q$  when  $q \rightarrow 1$  so we obtain :

$$\lim_{q \rightarrow 1} \chi_{ijk}(q) = 3 - \Lambda_{ij}(1, 1) - \Lambda_{ik}(1, 1) - \Lambda_{jk}(1, 1) + \Lambda_{ijk}(1, 1, 1)$$

Now, let's develop the other term. Using the equation (16) in (Engelke and Hitz 2020), if we note  $\lambda$  the density of the exponent measure with respect to the Lebesgue measure, we have :

$$\mathbb{P}(Y_i > 1, Y_j > 1 \mid Y_k > 1) = \int_{[1, \infty[^3} \lambda(y) dy = \Lambda([1, \infty[^3)$$

The properties of valid density give us 3 equations that we can use to get that :

$$3 = \Lambda([0, \infty[ \times [1, \infty[^2) + \Lambda([1, \infty[ \times ]0, \infty[ \times [1, \infty[) + \Lambda([1, \infty[^2 \times ]0, \infty[)$$

These quantities can be decomposed as several “volume” in the space. We can merge some volume between them to obtain this relation :

$$3 = \Lambda_{ij}(1, 1) + \Lambda_{ik}(1, 1) + \Lambda_{jk}(1, 1) - \Lambda_{ijk}(1, 1, 1) + \Lambda([1, \infty[^3)$$

and we have the right formula by isolating the  $\Lambda([1, \infty[^3)$  term.

### 3.2 Inequalities of section 2

- $\chi_{ij}^k \leq |A| \chi_{ij}^A$  :

Let  $k \in A$ . We have by definition that :

$$\begin{aligned} \mathbb{P}(F(X_i) > q \mid F(X_j) > q, F(X_k) > q) &= \frac{\mathbb{P}(F(X_i) > q, F(X_j) > q, F(X_k) > q)}{\mathbb{P}(F(X_j) > q, F(X_k) > q)} \\ &\leq \frac{\mathbb{P}(F(X_i) > q, F(X_j) > q, \|F(X_A)\|_\infty > q)}{\mathbb{P}(F(X_j) > q, F(X_k) > q)} \\ &\leq \chi_{ij}^A(q) \frac{\mathbb{P}(F(X_j) > q, \|F(X_A)\|_\infty > q)}{\mathbb{P}(F(X_j) > q, F(X_k) > q)} \\ &\leq \chi_{ij}^A(q) \sum_{k \in A} \frac{\mathbb{P}(F(X_j) > q, F(X_k) > q)}{\mathbb{P}(F(X_j) > q, F(X_k) > q)} \\ &\leq |A| \times \chi_{ij}^A(q) \end{aligned}$$

If we take the limit of both side, we get :

$$\chi_{ij}^k \leq |A| \times \chi_{ij}^A$$

which shows the first inequality.

- $\chi_{ij}^A \leq \sum_{k \in A} \chi_{ij}^k$  :

We use the alternative writing for this proof. If we apply again the definition, we have :

$$\begin{aligned} \chi_{ij}^A &= \frac{\mathbb{P}(Y_i > 1, Y_j > 1, \|Y_A\|_\infty > 1)}{\mathbb{P}(Y_j > 1, \|Y_A\|_\infty > 1)} \\ &\leq \sum_{k \in A} \frac{\mathbb{P}(Y_i > 1, Y_j > 1, Y_k > 1)}{\mathbb{P}(Y_j > 1, \|Y_A\|_\infty > 1)} \\ &\leq \sum_{k \in A} \chi_{ij}^k \frac{\mathbb{P}(Y_j > 1, Y_k > 1)}{\mathbb{P}(Y_j > 1, \|Y_A\|_\infty > 1)} \end{aligned}$$

but as  $\mathbb{P}(Y_j > 1, Y_k > 1) \leq \mathbb{P}(Y_j > 1, \|Y_A\|_\infty > 1)$  for all  $k \in A$ , we finally have :

$$\chi_{ij}^A \leq \sum_{k \in A} \chi_{ij}^k$$

which is what we wanted.

## References

- Engelke, Sebastian, and Adrien S. Hitz. 2020. “Graphical Models for Extremes.” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 82 (4): 871–932. <https://doi.org/10.1111/rssb.12355>.
- Ferreira, H. 2011. “Dependence Between Two Multivariate Extremes.” *Statistics & Probability Letters* 81 (5): 586–91. <https://doi.org/10.1016/j.spl.2011.01.014>.
- Hentschel, Manuel, Sebastian Engelke, and Johan Segers. 2023. “Statistical Inference for Hüsler-Reiss Graphical Models Through Matrix Completions.” arXiv. <https://doi.org/10.48550/arXiv.2210.14292>.
- Strokorb, Kirstin. 2020. “Extremal Independence Old and New.” arXiv. <https://doi.org/10.48550/arXiv.2002.07808>.