.

WISHART/RIETZ DISTRIBUTIONS AND DECOMPOSABLE UNDIRECTED GRAPHS

BY

STEEN ANDERSSON

INDIANA UNIVERSITY

AND

THOMAS KLEIN

AUGSBURG UNIVERSITÄET

TARTU 26JUN2007

Let V be a finite set and let $\mathbf{PD}(V)$ denote the open convex cone of all positive definite $V \times V$ matrices. Let $\Sigma \in \mathbf{PD}(V)$ and $\lambda > \frac{V-1}{2}$*. The classical Wishart distribution $\mathbb{W}_{\Sigma,\lambda}$ on $\mathbf{PD}(V)$ with *shape parameter* λ and *expectation* Σ is defined by

$$(0.1)\quad d\mathbb{W}_{\Sigma,\lambda}(S) := \frac{\pi^{\frac{V(1-V)}{4}} \lambda^{\lambda V} |S|^{\lambda - \frac{V+1}{2}}}{\prod(\Gamma(\lambda - \frac{i-1}{2}) | i = 1, \cdots, V) |\Sigma|^{\lambda}} \exp\{-\lambda \mathrm{Tr}(\Sigma^{-1} S)\} dS.$$

The parameter Σ deserves its name since the expectation $\mathbb{E}(W_{\Sigma,\lambda}) = \Sigma$.

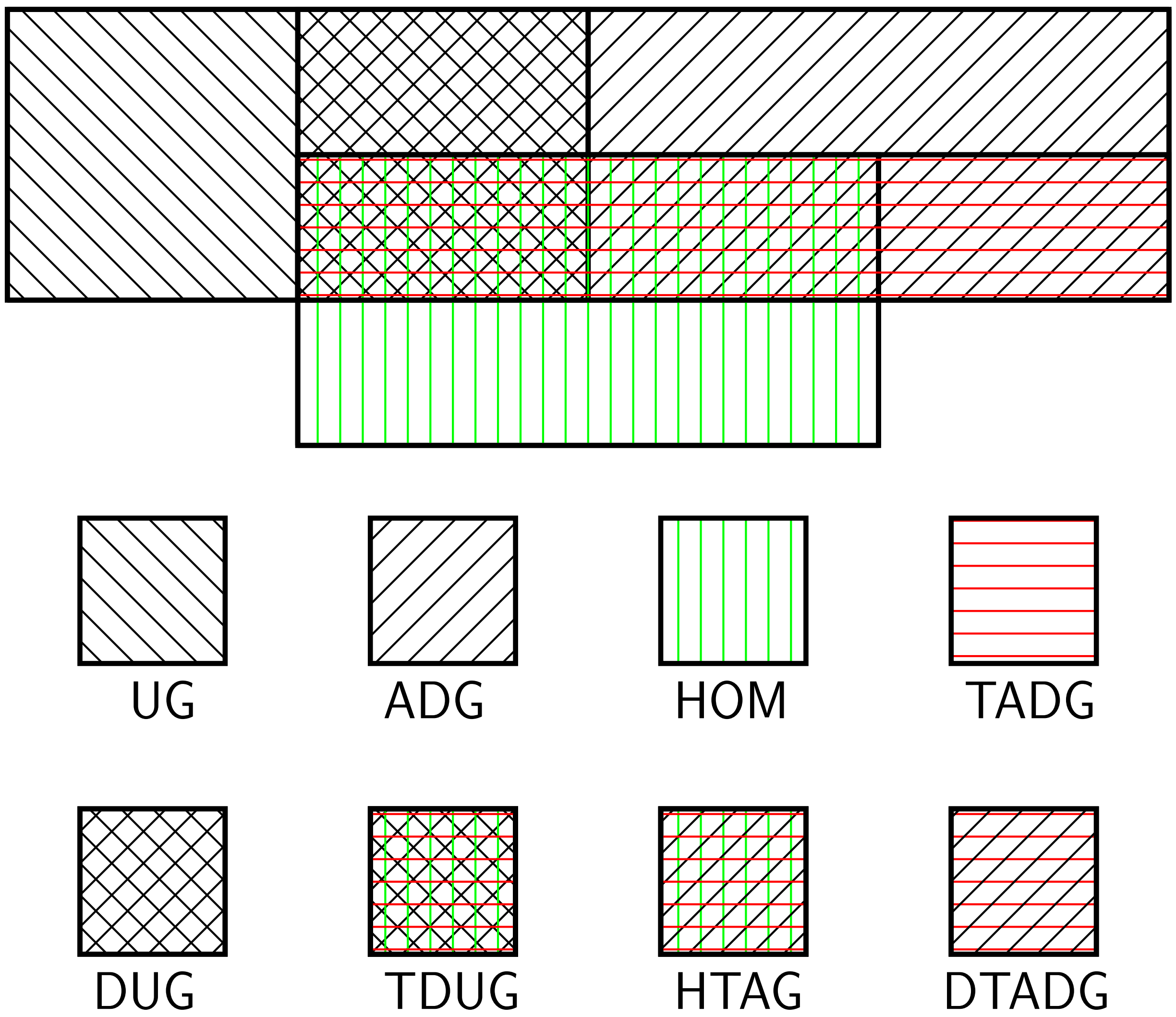
The statistical model $(W_{\Sigma,\lambda} \in \mathcal{P}(\mathrm{S}(V)) | \Sigma \in \mathbf{PD}(V))$, where $\mathcal{P}(\mathrm{S}(V))$ denotes the set of probability measures on the vector space $\mathrm{S}(V)$ of symmetric $V \times V$ matrices, is well-known to be a full regular exponential family in its expectation parametrization.

The corresponding *natural parameter* is then $\Delta := \lambda \Sigma^{-1} \in \mathbf{PD}(V)$ and with this parameter (0.1), in this parametrization denoted $W_{\Delta,\lambda}$, takes the form

$$(0.2)\qquad dW_{\Delta,\lambda}(S) := \frac{\pi^{\frac{V(1-V)}{4}} |\Delta|^{\lambda} |S|^{\lambda - \frac{V+1}{2}}}{\prod(\Gamma(\lambda - \frac{i-1}{2}) | i = 1, \cdots, V)} \exp\{-\mathrm{Tr}(\Delta S)\} dS.$$

*The notation V of a finite set also denote the cardinality of the set

Context and motivation.



UG: Undirected Graphs, ADG: Acyclic directed graphs, HOM: Homogeneous cones, TADG: **Transitive** ADG, DUG: Decomposable UG, TDUG: DUG without four chains, HTADG: TADG with transitive action, and DTADG: TADG with a least one diamond pattern.

Notation

Let $v_1, \cdots, v_V$ be a numbering of the elements of V. As it is customary the set V is then identified with its numbering, i.e., v_i is just denoted by i, $i = 1, \cdots, V$.

Set $< i > := \{1, \cdots, i-1\}$. Note that $< 1 > = \emptyset$ and that $< 2 > = \{1\}$.

For $\Sigma \in \mathbf{PD}(V)$ and $i = 1, \cdots, V$, let $\Sigma_{[i]}$, $\Sigma_{<i>}$, $\Sigma_{[i>}$, and $\Sigma_{<i]}$ denote the $\{i\} \times \{i\}$, $< i > \times < i >$, the $\{i\} \times < i >$, and the $< i > \times \{i\}$ sub-matrices of $\Sigma \in \mathbf{PD}(V)$, respectively, and define

$$\Sigma_{[i]\bullet} = \Sigma_{[i]} - \Sigma_{[i>}(\Sigma_{<i>})^{-1}\Sigma_{<i]} > 0.$$

The rational functions $\Sigma_{[i]\bullet}$ of the entries of the matrix Σ can also be defined through the unique decomposition $\Sigma = TDT^t$, where T is a lower triangular matrix with all diagonal elements equal to 1 and D is a diagonal matrix with positive diagonal elements, since $D_{ii} = \Sigma_{[i]\bullet}$, $i = 1, \cdots, V$.

Thus corresponding to the opposite ordering of the elements of V we have the unique decomposition $\Delta = T^t DT$, $\Delta \in \mathbf{PD}(V)$, where again T is a lower triangular matrix with all diagonal elements equal to 1 and D is a diagonal matrix with positive diagonal elements. We can then similarly define $\Delta_{[i]\circ} := D_{ii}$, $i = 1, \cdots, V$.

Matrix Picture.

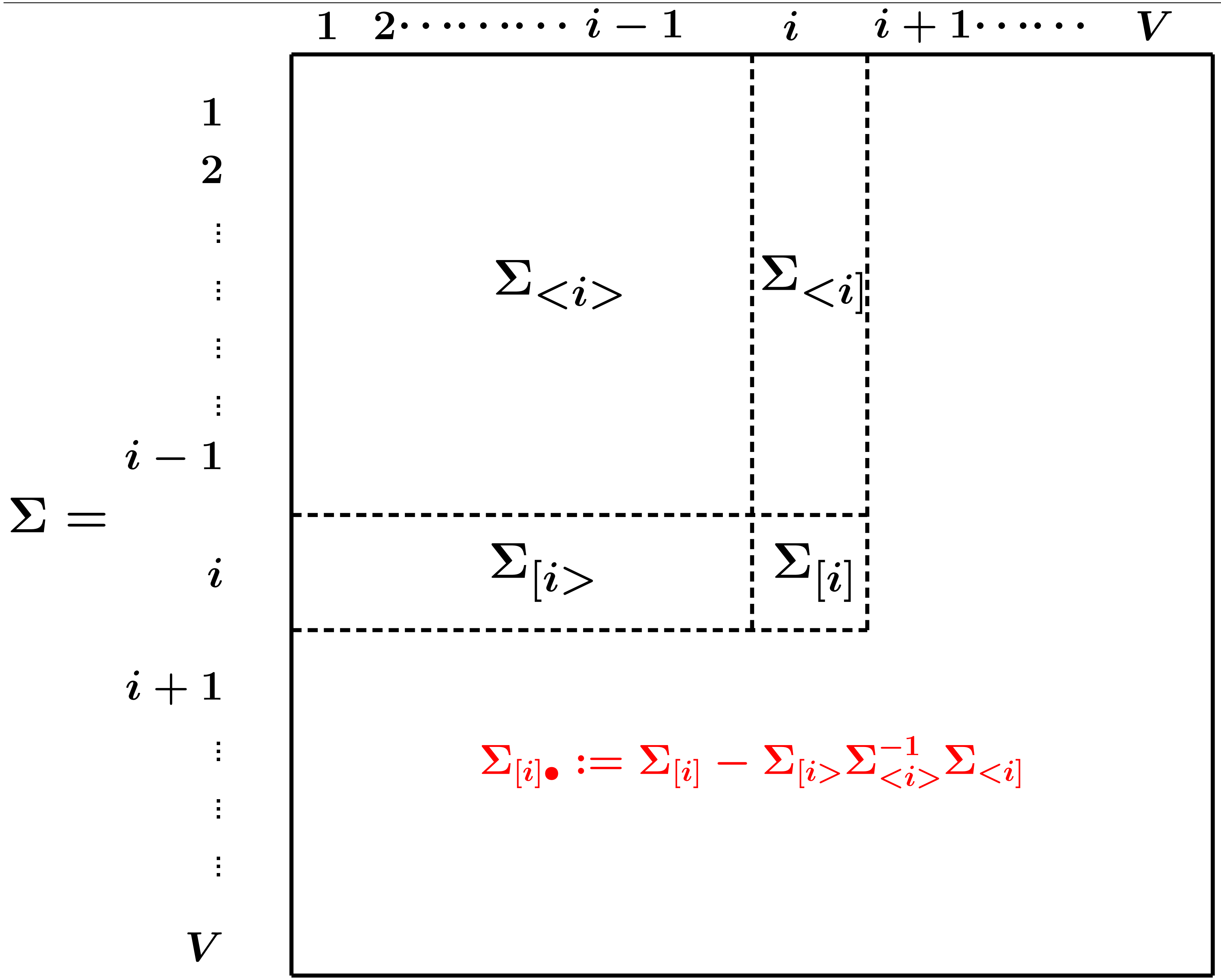


Illustration of decomposition.

Illustration of $\boldsymbol{\Sigma} = \boldsymbol{T}\boldsymbol{D}\boldsymbol{T}^{\boldsymbol{t}}$ in a case where $\boldsymbol{V} = \boldsymbol{8}$: .

$$
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & 1 & 0 & 0 & 0 & 0 \\
 & & & & 1 & 0 & 0 & 0 \\
 & & & & & 1 & 0 & 0 \\
 & & & & & & 1 & 0 \\
 & & & & & & & 1
\end{bmatrix}
\cdot
\begin{bmatrix}
+ & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & + & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & + & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & + & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & + & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & + & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & + & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & +
\end{bmatrix}
\cdot
\begin{bmatrix}
1 & & & & & & & \\
0 & 1 & & & & & & \\
0 & 0 & 1 & & & & & \\
0 & 0 & 0 & 1 & & & & \\
0 & 0 & 0 & 0 & 1 & & & \\
0 & 0 & 0 & 0 & 0 & 1 & & \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
$$

Rietz integral.

The classical *Rietz integral*, defined for any $\boldsymbol{\lambda} \equiv (\boldsymbol{\lambda}_i | \boldsymbol{i} = \mathbf{1}, \cdots, \boldsymbol{V}) \in \mathbb{R}^{\boldsymbol{V}}$, is†

$$\int_{\mathrm{PD}(\boldsymbol{V})} \prod((\boldsymbol{S}_{[i]\bullet})^{\lambda_i - \frac{V+1}{2}} | \boldsymbol{i} = \mathbf{1}, \cdots, \boldsymbol{V}) \exp\{-\mathrm{Tr}(\boldsymbol{\Delta S})\} \boldsymbol{dS},$$

convergent if and only if

$$\boldsymbol{\lambda}_i > \frac{\boldsymbol{i} - \mathbf{1}}{\mathbf{2}}, \quad \boldsymbol{i} = \mathbf{1}, \cdots, \boldsymbol{V},$$

and in this case with the value

$$\boldsymbol{\pi}^{\frac{V(V-1)}{4}} \prod (\frac{\Gamma(\boldsymbol{\lambda}_i - \frac{i-1}{2})}{(\boldsymbol{\Delta}_{[i]\circ})^{\lambda_i}} | \boldsymbol{i} = \mathbf{1}, \cdots, \boldsymbol{V}).$$

†In fact only defined for $\boldsymbol{\Delta} = \mathbf{1}_{\boldsymbol{V}}$.

We can thus define the distributions on $\mathbf{PD}(V)$ as follows:

(0.3)

$$dR_{\Delta,\lambda}(S) := \pi^{\frac{V(1-V)}{4}} \prod \left(\frac{(\Delta_{[i]\circ})^{\lambda_i} (S_{[i]\bullet})^{\lambda_i - \frac{V+1}{2}}}{\Gamma(\lambda_i - \frac{i-1}{2})} \middle| i = 1, \cdots, V \right) \exp\{-\mathrm{Tr}(\Delta S)\} dS.$$

Definition 0.1. The probability $R_{\Delta,\lambda}$ is called the *classical Rietz distribution* on $\mathbf{PD}(V)$ wrt. the ordering $1, \cdots, V$ of the elements of V and with *shape parameter* $\lambda \equiv (\lambda_i | i = 1, \cdots, V)$ and *natural parameter* Δ.

Note when λ_i does not depend on $i \in V$ we regain the Wishart distribution above.

For the numbering $\mathbf{1}, \cdots, \boldsymbol{V}$ of the elements of $\boldsymbol{V}$ we can for any $\boldsymbol{\Sigma} \in \mathbf{PD}(\boldsymbol{V})$ and any $\boldsymbol{\lambda} \equiv (\boldsymbol{\lambda}_i | \boldsymbol{i} = \mathbf{1}, \cdots, \boldsymbol{V}) \in \mathbb{R}_+^{\boldsymbol{V}}$ define
$\boldsymbol{\Sigma}^{-\boldsymbol{\lambda}} := (\boldsymbol{T}^t)^{-1}\mathbf{Diag}(\frac{\lambda_i}{\Sigma_{[i]\bullet}} | \boldsymbol{i} = \mathbf{1}, \cdots, \boldsymbol{V})\boldsymbol{T}^{-1}$,
where $\boldsymbol{\Sigma} = \boldsymbol{T}\mathbf{Diag}(\boldsymbol{\Sigma}_{[i]\bullet} | \boldsymbol{i} = \mathbf{1}, \cdots, \boldsymbol{V})\boldsymbol{T}^t$ with $\boldsymbol{T}$ being lower triangular with 1's in the diagonal.
In the case where $\boldsymbol{\lambda}_i = \boldsymbol{\mu}$, independent of $\boldsymbol{i} = \mathbf{1}, \cdots, \boldsymbol{V}$ we get $\boldsymbol{\Sigma}^{-\boldsymbol{\lambda}} = \boldsymbol{\mu}\boldsymbol{\Sigma}^{-1}$. Note that $\boldsymbol{\Sigma}^{-\boldsymbol{\lambda}} = \boldsymbol{\Delta} \in \mathbf{PD}(\boldsymbol{V})$ has a unique solution in $\boldsymbol{\Sigma} \in \mathbf{PD}(\boldsymbol{V})$ denoted $\boldsymbol{\Sigma} = \boldsymbol{\Delta}^{\boldsymbol{\lambda}-}$ (opposite ordering).

By a relative simple calculation or as a special case of a later result it follows that
$\mathbb{E}(\boldsymbol{R}_{\boldsymbol{\Delta},\boldsymbol{\lambda}}) = \boldsymbol{\Sigma} := \boldsymbol{\Delta}^{-\boldsymbol{\lambda}}$.
Thus if the natural parameter $\boldsymbol{\Delta} \in \mathbf{PD}(\boldsymbol{V})$ is replaced by the *expectation parameter* $\boldsymbol{\Sigma} \in \mathbf{PD}(\boldsymbol{V})$ we obtain

(0.4)
$$\boldsymbol{d}\mathbb{R}_{\boldsymbol{\Sigma},\boldsymbol{\lambda}}(\boldsymbol{S}) := \boldsymbol{\pi}^{\frac{V(1-V)}{4}} \prod \left(\frac{\boldsymbol{\lambda}_i^{\boldsymbol{\lambda}_i}(\boldsymbol{S}_{[i]\bullet})^{\boldsymbol{\lambda}_i - \frac{V+1}{2}}}{\boldsymbol{\Gamma}(\boldsymbol{\lambda}_i - \frac{i-1}{2})(\boldsymbol{\Sigma}_{[i]\bullet})^{\boldsymbol{\lambda}_i}} \middle| \boldsymbol{i} = \mathbf{1}, \cdots, \boldsymbol{V}\right) \mathbf{exp}\{-\mathbf{Tr}(\boldsymbol{\Sigma}^{-\boldsymbol{\lambda}}\boldsymbol{S})\}\boldsymbol{dS}.$$

Acyclic Mixed Graphs (AMG)

Let $\mathcal{V} \equiv (V, E)$ be an acyclic mixed graph (AMG). The equivalence relation $\sim$ on the vertex set V given by $v_1 \sim v_2$ iff $v_1 = v_2$ or v_1 and v_2 are connected by an undirected path, defines the set of equivalence classes $V/\sim$ called the set of *boxes* of the AMG. Each box $B \in V/\sim$ thus corresponds to a subset, denoted $[B]$, of V and the by $[B]$ induced graph is an undirected graph, called the $[B]$ box graph. We shall assume that the AMG has two important properties

(1) All box graphs are complete, and

(2) The AMG has no triplexes. (i.e., no flags $\bullet - \bullet \leftarrow \bullet$ and no immoralities $\bullet \rightarrow \bullet \leftarrow \bullet$)

The *skeleton* of $\mathcal{V}$ is given by $\mathcal{U}(\mathcal{V}) \equiv \mathcal{U} := (V, E \cup E^o)$, where $(v_1, v_2)^o := (v_2, v_1)$, $(v_1, v_2) \in V^2$. Since $\mathcal{V}$ has no triplexes $\mathcal{U}$ is decomposable.
Let $B \in V/\sim$ and let $v \in [B]$. The *parents* of v, the subset
$\mathrm{pa}_{\mathcal{V}}(v) \equiv \mathrm{pa}(v) := \{v' \in V | v' \rightarrow_{\mathcal{V}} v\}$,
is independent of $v \in [B]$ and thus we can define $< B > := \mathrm{pa}(v)$, $v \in [B]$.

The AMP Markov properties for $\mathcal{V}$ are equivalent to the Markov properties of its skeleton $\mathcal{U}$.

Consider centered regular normal distributions $N(0, \Sigma)$ on $\mathbb{R}^V$, $\Sigma \in \mathbf{PD}(V)$. Thus the Markov properties given by the AMG $\mathcal{V}$ or equivalently $\mathcal{U}$ impose restrictions on the variance matrix Σ or equivalently on the precision matrix $\Delta := \Sigma^{-1}$. For our special type of AMG $\mathcal{V}$ these restriction are easily expressed in terms of Δ:

Variance and precision.

Let $\mathcal{U} \equiv (V, F)$ be an undirected graph with vertex set V and let $\mathrm{S}(V)$ denote the vector space of symmetric $V \times V$ matrices $S \equiv (S_{uv} | (u, v) \in V^{\times 2})$.

$$\mathrm{S}(\mathcal{U}) := \{S \in \mathrm{S}(V) | \forall u, v \in V : \text{ with } u \neq v \text{ and } (u, v) \notin F : S_{uv} = 0\}.$$

Thus we have the *projection mapping* $p_{\mathcal{U}} \equiv p$

$$\begin{aligned} p : \mathrm{S}(V) &\longrightarrow \mathrm{S}(\mathcal{U}) \\ S &\longrightarrow p(S), \end{aligned}$$

where

$$p(S)_{uv} := \begin{cases} S_{uv} & \text{if} \quad (u, v) \in F \quad \text{or} \quad u = v \\ 0 & \text{if} \quad (u, v) \notin F \quad \text{and} \quad u \neq v \end{cases}.$$

Set $\mathrm{PD}^0(\mathcal{U}) := \mathrm{S}(\mathcal{U}) \cap \mathrm{PD}(V)$. It is well-known that $N(0, \Sigma)$ satisfies the Markov properties given by $\mathcal{U}$ if and only if $\Delta = \Sigma^{-1} \in \mathrm{PD}^0(\mathcal{U})$. Set $\mathrm{PD}(\mathcal{U}) := \mathrm{PD}^0(\mathcal{U})^{-1}$. Then this subset $\mathrm{PD}(\mathcal{U}) \subseteq \mathrm{PD}(V)$ can similarly be described by $\Sigma \in \mathrm{PD}(\mathcal{U})$ if and only if $N(0, \Sigma)$ satisfies the Markov properties given by $\mathcal{U}$.

Also set

$$\mathrm{P}(\mathcal{U}) := \{S \in \mathrm{S}(\mathcal{U}) | S_C \in \mathrm{PD}(C) \;\; \forall \;\text{ cliques } C \subset V\}$$

where S_C denote the $C \times C$ submatrix of S.

Proposition 0.1. *The mapping*

$$
\begin{aligned}
\mathbf{PD}(\boldsymbol{\mathcal{U}}) &\to \mathbf{P}(\boldsymbol{\mathcal{U}}) \\
\boldsymbol{S} &\to \boldsymbol{p}(\boldsymbol{S}) \\
\hat{\boldsymbol{S}} &\leftarrow \boldsymbol{S}
\end{aligned}
\tag{0.5}
$$

is a well defined one-to-one correspondence.

The open convex cones $\mathbf{P}(\boldsymbol{\mathcal{U}})$ *and* $\mathbf{PD}^0(\boldsymbol{\mathcal{U}})$ *are dual to each other through the isomorphism (of open convex cones)*

$$
\begin{aligned}
\mathbf{P}(\boldsymbol{\mathcal{U}}) &\leftrightarrow (\mathbf{PD}^0(\boldsymbol{\mathcal{U}}))^* \\
\boldsymbol{S} &\to (\boldsymbol{T} \to \mathbf{Tr}(\boldsymbol{ST})) \\
(\boldsymbol{S} \to \mathbf{Tr}(\boldsymbol{TS})) &\leftarrow \boldsymbol{T}.
\end{aligned}
$$

Proof. The two results are well known and proved by induction after the number of cliques in $\boldsymbol{\mathcal{U}}$. □

The standard inverse mappings

The inverse mappings

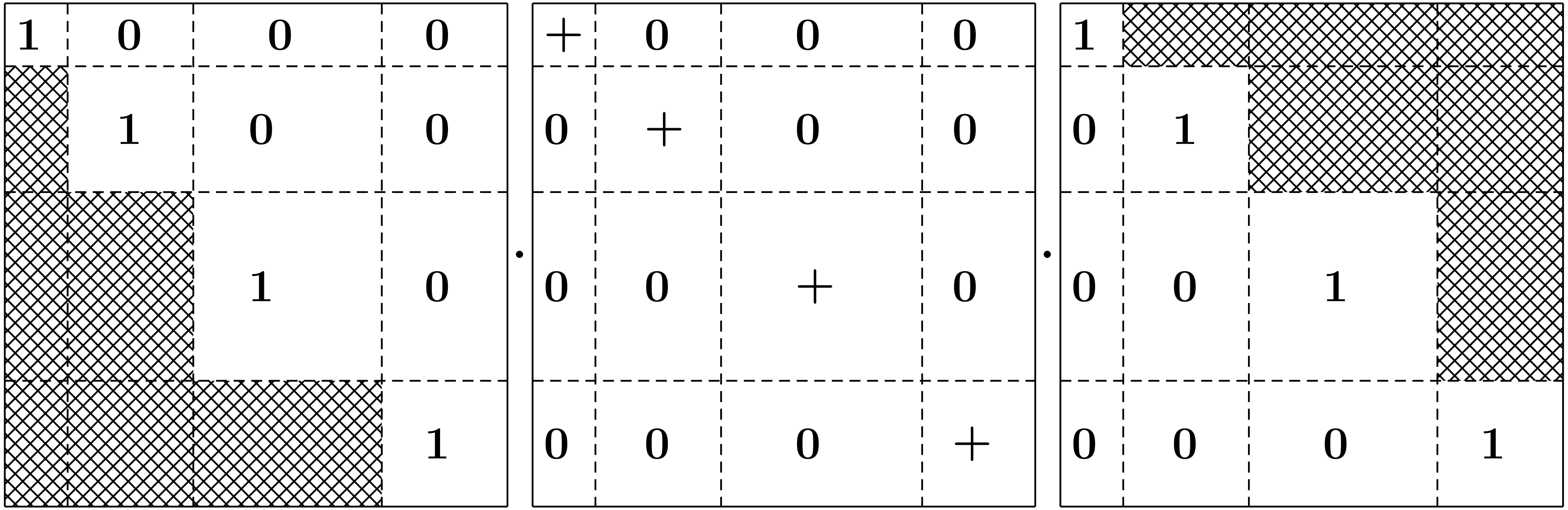
$$\begin{aligned} \mathrm{P}(\boldsymbol{\mathcal{U}}) &\leftrightarrow \mathrm{PD}^0(\boldsymbol{\mathcal{U}}) \\ \boldsymbol{S} &\rightarrow (\hat{\boldsymbol{S}})^{-1} =: \boldsymbol{S}^{-1} \\ \boldsymbol{p}(\boldsymbol{T}^{-1}) &\leftarrow \boldsymbol{T}. \end{aligned}$$

defines a one-to-one correspondence.

Nevertheless, we shall make use of a special variation of this inverse mapping depending on the representation of $\boldsymbol{\mathcal{U}}$ as an AMG $\boldsymbol{\mathcal{V}}$ with the properties described above:
Let $\boldsymbol{\lambda} \equiv (\boldsymbol{\lambda}_B | \boldsymbol{B} \in \boldsymbol{V}/\sim) \in \mathbb{R}_+^{\boldsymbol{V}/\sim}$. For $\boldsymbol{S} \in \mathrm{P}(\boldsymbol{\mathcal{U}})$ the positive definite matrix $\hat{\boldsymbol{S}} \in \mathrm{PD}(\boldsymbol{\mathcal{U}})$ has a unique decomposition of the form $\hat{\boldsymbol{S}} = \boldsymbol{T}\boldsymbol{D}\boldsymbol{T}^t$, where $\boldsymbol{T}$ and $\boldsymbol{D}$ are block matrices according to the decomposition $\boldsymbol{V} = \dot{\cup}([\boldsymbol{B}] | \boldsymbol{B} \in \boldsymbol{V}/\sim)$, i.e., $\boldsymbol{T} \equiv (\boldsymbol{T}_{\boldsymbol{BB'}} | (\boldsymbol{B}, \boldsymbol{B'}) \in (\boldsymbol{V}/\sim)^{\times 2})$ and $\boldsymbol{D} \equiv (\boldsymbol{D}_{\boldsymbol{BB'}} | (\boldsymbol{B}, \boldsymbol{B'}) \in (\boldsymbol{V}/\sim)^{\times 2})$, satisfying that $\boldsymbol{D}_{\boldsymbol{BB}} \in \mathrm{PD}(\boldsymbol{B})$, $\boldsymbol{B} \in \boldsymbol{V}/\sim$, $\boldsymbol{D}_{\boldsymbol{BB'}} = \boldsymbol{0}$ for $\boldsymbol{B}, \boldsymbol{B'} \in \boldsymbol{V}/\sim$ with $\boldsymbol{B} \neq \boldsymbol{B'}$, i.e., $\boldsymbol{D} \equiv \mathrm{Diag}(\boldsymbol{D}_{\boldsymbol{BB}} | \boldsymbol{B} \in \boldsymbol{V}/\sim)$ is block diagonal with positive definite entries in the diagonal, $\boldsymbol{T}_{\boldsymbol{BB}} = \boldsymbol{1}_{\boldsymbol{B}}$, $\boldsymbol{B} \in \boldsymbol{V}/\sim$, $\boldsymbol{T}_{\boldsymbol{BB'}} = \boldsymbol{0}$, $\boldsymbol{B}, \boldsymbol{B'} \in \boldsymbol{V}/\sim$ when there is not partly directed path from $\boldsymbol{B}$ to $\boldsymbol{B'}$, i.e., $\boldsymbol{T}$ is "lower block triangular" with identity matrices in the diagonal and possible extra zeroes.

Illustration of the decomposition.

Illustration of $\hat{\boldsymbol{\Sigma}} = \boldsymbol{T}\boldsymbol{D}\boldsymbol{T}^t$ in a case where $\boldsymbol{V}/\sim = 4$: .



The $\boldsymbol{\lambda}$-inverse mappings

$$\begin{aligned} \mathrm{P}(\mathcal{U}) &\leftrightarrow \mathrm{PD}^0(\mathcal{U}) \\ S = p(T\mathrm{Diag}(D_{BB}|B \in V/\sim)T^t) &\longrightarrow S^{-\lambda} := \\ & \quad (T^t)^{-1}\mathrm{Diag}(\lambda_B D_{BB}^{-1}|B \in V/\sim)T^{-1}) \\ p(U^{-1}\mathrm{Diag}(\lambda_B E_{BB}^{-1}|B \in V/\sim)(U^t)^{-1}) &\longleftarrow D = U^t\mathrm{Diag}(E_{BB}|B \in V/\sim)U \\ =: D^{\lambda-} & \end{aligned}$$

For $\Sigma \in \mathrm{P}(\mathcal{U})$ we denote the diagonal matrices D_{BB} of D in the unique decomposition $\Sigma = p(TDT)$ given above by $\Sigma_{[B]\bullet} = D_{BB}$, $B \in V/\sim$. Analogous for $\Delta \in \mathrm{PD}^0(\mathcal{U})$ we set $\Delta_{[B]\circ} := D_{BB}$, $B \in V/\sim$, when $\Delta = T^tDT$. thus we have

$$\begin{aligned} (\Sigma^{-\lambda})_{[B]\circ} &= \lambda_B(\Sigma_{[B]\bullet})^{-1} \\ (\Delta^{\lambda-})_{[B]\bullet} &= \lambda_B(\Delta_{[B]\circ})^{-1} \end{aligned}, \quad B \in V/\sim .$$

Let $\boldsymbol{\mathcal{V}}$ be the AMG from above, i.e., having the properties (1) and (2), and let $\boldsymbol{\mathcal{U}}$ be the underlying skeleton.

For $\boldsymbol{S} \in \mathrm{S}(\boldsymbol{V})$ define $\boldsymbol{S}_{[\boldsymbol{B}]}$, $\boldsymbol{S}_{[\boldsymbol{B}>}$, and $\boldsymbol{S}_{<\boldsymbol{B}>}$ as the $[\boldsymbol{B}] \times [\boldsymbol{B}]$, $[\boldsymbol{B}]\times < \boldsymbol{B} >$, and $< \boldsymbol{B} > \times < \boldsymbol{B} >$ sub-matrices of $\boldsymbol{S}$.

Let $\boldsymbol{M} \in \boldsymbol{V}/\sim$ be a maximal box and let $\boldsymbol{\mathcal{V}}_{\boldsymbol{M}}$ be the AMG induced by the subset $\boldsymbol{V}_{\boldsymbol{M}} := \boldsymbol{V}\backslash[\boldsymbol{M}]$. Then $\boldsymbol{V}_{\boldsymbol{M}}/\sim = (\boldsymbol{V}/\sim)\backslash\{\boldsymbol{M}\}$.

Furthermore $[\boldsymbol{B}]$, $< \boldsymbol{B} >$, $\boldsymbol{S}_{[\boldsymbol{B}]}$, $\boldsymbol{S}_{[\boldsymbol{B}>}$, and $\boldsymbol{S}_{<\boldsymbol{B}>}$, $\boldsymbol{B} \in \boldsymbol{V}_{\boldsymbol{M}}/\sim$ are the same in $\boldsymbol{\mathcal{V}}$ and $\boldsymbol{\mathcal{V}}_{\boldsymbol{M}}$.

The skeleton $\boldsymbol{\mathcal{U}}_{\boldsymbol{M}}$ of $\boldsymbol{\mathcal{V}}_{\boldsymbol{M}}$ is the subgraph of $\boldsymbol{\mathcal{U}}$ induced by the subset $\boldsymbol{V}_{\boldsymbol{M}}$ of $\boldsymbol{V}$.

Proposition 0.2. *The mapping*

$$\mathrm{PD}([\boldsymbol{M}]) \times \mathbb{R}^{[M]\times<M>} \times \mathrm{P}(\boldsymbol{\mathcal{U}}_M) \to \mathrm{P}(\boldsymbol{\mathcal{U}})$$
$$(\boldsymbol{L}_M, \boldsymbol{R}_M, \boldsymbol{S}_M) \to$$

$$\boldsymbol{p}\left(\begin{pmatrix} \mathbf{1}_{V_M} & \mathbf{0} \\ \boldsymbol{R}_{M\mathbf{0}} & \mathbf{1}_{[M]} \end{pmatrix}\begin{pmatrix} \hat{\boldsymbol{S}}_M & \mathbf{0} \\ \mathbf{0} & \boldsymbol{L}_M \end{pmatrix}\begin{pmatrix} \mathbf{1}_{V_M} & \boldsymbol{R}^t_{M\mathbf{0}} \\ \mathbf{0} & \mathbf{1}_{[M]} \end{pmatrix}\right) = \tag{0.6}$$

$$\boldsymbol{p}\left(\begin{pmatrix} \hat{\boldsymbol{S}}_M & \hat{\boldsymbol{S}}_M\boldsymbol{R}^t_{M\mathbf{0}} \\ \boldsymbol{R}_{M\mathbf{0}}\hat{\boldsymbol{S}}_M & \boldsymbol{L}_M + \boldsymbol{R}_{M\mathbf{0}}\hat{\boldsymbol{S}}_M\boldsymbol{R}^t_{M\mathbf{0}} \end{pmatrix}\right)$$

where the $[\boldsymbol{M}] \times \boldsymbol{V}_M$ *matrix* $\boldsymbol{R}_{M\mathbf{0}}$ *is given by* $(\boldsymbol{R}_{M\mathbf{0}})_{[M]\times<M>} = \boldsymbol{R}_M$ *and* $(\boldsymbol{R}_{M\mathbf{0}})_{[M]\times(V_M\setminus<M>)} = \mathbf{0}_{[M]\times(V_M\setminus<M>)}$*, is well-defined and a bijection.*

Transformation Proposition Continued.

The inverse mapping takes the form

$$\begin{aligned} \mathbf{PD}([\boldsymbol{M}]) \times \mathbb{R}^{[M]\times<M>} \times \mathbf{P}(\boldsymbol{\mathcal{U}}_M) &\leftarrow \mathbf{P}(\boldsymbol{\mathcal{U}}) \\ (\hat{\boldsymbol{S}}_{[M]\bullet}, \hat{\boldsymbol{S}}_{[M>\bullet}, \boldsymbol{S}_{\mathbf{V}_M\times\mathbf{V}_M}) &\leftarrow \boldsymbol{S} \end{aligned} \tag{0.7}$$

where $\hat{\boldsymbol{S}}_{[M]\bullet} := \boldsymbol{S}_{[M]} - \boldsymbol{S}_{[M>}\hat{\boldsymbol{S}}^{-1}_{<M>}\boldsymbol{S}_{<M]}$ and $\hat{\boldsymbol{S}}_{[M>\bullet} := \boldsymbol{S}_{[M>}\hat{\boldsymbol{S}}^{-1}_{<M>}$.

Note that $\hat{\boldsymbol{S}}^{-1}_{<M>} := ((\hat{\boldsymbol{S}})_{<M>})^{-1}$.

Furthermore the "hats" can be taken away since the skeleton of the subgraph induced by $<\boldsymbol{M}>$ and $[\boldsymbol{M}]$ are complete (no triplexes), i.e.,

$$\boldsymbol{S}_{[M]\bullet} := \boldsymbol{S}_{[M]} - \boldsymbol{S}_{[M>}\boldsymbol{S}^{-1}_{<M>}\boldsymbol{S}_{<M]} \ \ \textbf{and} \ \ \boldsymbol{S}_{[M>\bullet} := \boldsymbol{S}_{[M>}\boldsymbol{S}^{-1}_{<M>}.$$

Proof. The inverse of

$$\begin{pmatrix} \mathbf{1}_{V_M} & \mathbf{0} \\ \boldsymbol{R}_{M0} & \mathbf{1}_{[M]} \end{pmatrix} \begin{pmatrix} \hat{\boldsymbol{S}}_M & \mathbf{0} \\ \mathbf{0} & \boldsymbol{L}_M \end{pmatrix} \begin{pmatrix} \mathbf{1}_{V_M} & \boldsymbol{R}^t_{M0} \\ \mathbf{0} & \mathbf{1}_{[M]} \end{pmatrix}$$

is

$$\boldsymbol{\Delta} := \begin{pmatrix} \mathbf{1}_{V_M} & -\boldsymbol{R}^t_{M0} \\ \mathbf{0} & \mathbf{1}_{[M]} \end{pmatrix} \begin{pmatrix} \hat{\boldsymbol{S}}^{-1}_M & \mathbf{0} \\ \mathbf{0} & \boldsymbol{L}^{-1}_M \end{pmatrix} \begin{pmatrix} \mathbf{1}_{V_M} & \mathbf{0} \\ -\boldsymbol{R}_{M0} & \mathbf{1}_{[M]} \end{pmatrix} =$$

$$\begin{pmatrix} \hat{\boldsymbol{S}}^{-1}_M + \boldsymbol{R}^t_{M0}\boldsymbol{L}^{-1}_M\boldsymbol{R}_{M0} & -\boldsymbol{R}^t_{M0}\boldsymbol{L}^{-1}_M \\ -\boldsymbol{L}^{-1}_M\boldsymbol{R}_{M0} & \boldsymbol{L}^{-1}_M \end{pmatrix}.$$

Let $(\boldsymbol{u}, \boldsymbol{v}) \notin \boldsymbol{E} \cup \boldsymbol{E}^{\boldsymbol{o}}$ with $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{V}_M$. Then either $\boldsymbol{u} \notin < \boldsymbol{M} >$ or $\boldsymbol{v} \notin < \boldsymbol{M} >$ since the AMG $\boldsymbol{\mathcal{V}}$ is without triplexes. Thus $\boldsymbol{\Delta}_{\boldsymbol{uv}} = (\hat{\boldsymbol{S}}^{-1}_M + \boldsymbol{R}^t_{M0}\boldsymbol{L}^{-1}_M\boldsymbol{R}_{M0})_{\boldsymbol{uv}} = (\hat{\boldsymbol{S}}^{-1}_M)_{\boldsymbol{uv}} + (\boldsymbol{R}^t_{M0}\boldsymbol{L}^{-1}_M\boldsymbol{R}_{M0})_{\boldsymbol{uv}}$. Since $\hat{\boldsymbol{S}}^{-1}_M \in \mathbf{PD}^0(\boldsymbol{\mathcal{U}}_M)$ we get $(\hat{\boldsymbol{S}}^{-1}_M)_{\boldsymbol{uv}} = \mathbf{0}$. Furthermore $(\boldsymbol{R}^t_{M0}\boldsymbol{L}^{-1}_M\boldsymbol{R}_{M0})_{\boldsymbol{uv}} = \mathbf{0}$ since either $\boldsymbol{u} \notin < \boldsymbol{M} >$ or $\boldsymbol{v} \notin < \boldsymbol{M} >$.

Proof, cont.

For $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{M}$ there is nothing to prove. Thus let $\boldsymbol{u} \in \boldsymbol{M}$ and $\boldsymbol{v} \in \boldsymbol{V_M}$ with $(\boldsymbol{u}, \boldsymbol{v}) \notin \boldsymbol{E} \cup \boldsymbol{E^o}$. Thus $\boldsymbol{v} \notin < \boldsymbol{M} >$. This implies that $(-\boldsymbol{L}_{\boldsymbol{M}}^{-1}\boldsymbol{R}_{\boldsymbol{M}\mathbf{0}})_{\boldsymbol{uv}} = \mathbf{0}$. Since thus $\boldsymbol{\Delta} \in \mathbf{S}(\boldsymbol{\mathcal{U}}) \cap \mathbf{P}(\boldsymbol{V}) = \mathbf{PD}^{\mathbf{0}}(\boldsymbol{\mathcal{U}})$, we have $\boldsymbol{\Delta}^{-1} \in \mathbf{PD}(\boldsymbol{\mathcal{U}})$ and thus $\boldsymbol{p}(\boldsymbol{\Delta}^{-1}) \in \mathbf{P}(\boldsymbol{\mathcal{U}})$. Thus the mapping is thus well-defined.

It is straight forward to establish that the mapping is injective. Thus we only have to establish that it is surjective. Let $\boldsymbol{S} \in \mathbf{P}(\boldsymbol{\mathcal{U}})$. Then $\hat{\boldsymbol{S}} \in \mathbf{PD}(\boldsymbol{\mathcal{U}})$. Define $\hat{\boldsymbol{S}}_{[\boldsymbol{B}]\bullet} := \boldsymbol{S}_{[\boldsymbol{B}]} - \boldsymbol{S}_{[\boldsymbol{B}>}\hat{\boldsymbol{S}}_{<\boldsymbol{B}>}^{-1}\boldsymbol{S}_{<\boldsymbol{B}]}$ and $\hat{\boldsymbol{S}}_{[\boldsymbol{B}>\bullet} := \boldsymbol{S}_{[\boldsymbol{B}>}\hat{\boldsymbol{S}}_{<\boldsymbol{B}>}^{-1}$, $\boldsymbol{B} \in \boldsymbol{V}/\sim$, noting that $\boldsymbol{S}_{[\boldsymbol{B}]} = (\hat{\boldsymbol{S}})_{[\boldsymbol{B}]}$, and $\boldsymbol{S}_{[\boldsymbol{B}>} = (\hat{\boldsymbol{S}})_{[\boldsymbol{B}>}$. Also let $\hat{\boldsymbol{S}}_{\boldsymbol{V_M}}$ denote the $\boldsymbol{V_M} \times \boldsymbol{V_M}$ sub-matrix of $\hat{\boldsymbol{S}}$. It is obvious that $\hat{\boldsymbol{S}}_{[\boldsymbol{B}]\bullet} \in \mathbf{PD}([\boldsymbol{B}])$ and that $\hat{\boldsymbol{S}}_{[\boldsymbol{B}>\bullet} \in \mathbb{R}^{[\boldsymbol{B}]\times<\boldsymbol{B}>}$, particular for $\boldsymbol{B} = \boldsymbol{M}$. Furthermore $\hat{\boldsymbol{S}}_{\boldsymbol{V_M}}^{-1} := (\hat{\boldsymbol{S}}_{\boldsymbol{V_M}})^{-1} = ((\hat{\boldsymbol{S}})^{-1})_{\boldsymbol{V_M}} - (\hat{\boldsymbol{S}}^{-1})_{\boldsymbol{V_M}\times[\boldsymbol{M}]}[(\hat{\boldsymbol{S}}^{-1})_{[\boldsymbol{M}]}]^{-1}(\hat{\boldsymbol{S}}^{-1})_{\boldsymbol{V_M}\times[\boldsymbol{M}]}$. Suppose $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{V_M}$ with $(\boldsymbol{u}, \boldsymbol{v}) \notin \boldsymbol{E} \cup \boldsymbol{E^o}$. Then again either $\boldsymbol{u} \notin < \boldsymbol{M} >$ or $\boldsymbol{v} \notin < \boldsymbol{M} >$.

Proof, cont.

Since $\hat{\boldsymbol{S}}^{-1} \in \mathbf{PD}^0(\boldsymbol{\mathcal{U}})$ we have $(\hat{\boldsymbol{S}}^{-1})_{\boldsymbol{uv}} = \mathbf{0}$. In particular we $((\hat{\boldsymbol{S}}^{-1})_{\boldsymbol{V_M}})_{\boldsymbol{uv}} = \mathbf{0}$ and $((\hat{\boldsymbol{S}}^{-1})_{\boldsymbol{V_M}\times[\boldsymbol{M}]}[(\hat{\boldsymbol{S}}^{-1})_{[\boldsymbol{M}]}]^{-1}(\hat{\boldsymbol{S}}^{-1})_{\boldsymbol{V_M}\times[\boldsymbol{M}]})_{\boldsymbol{uv}} = \mathbf{0}$. Thus $((\hat{\boldsymbol{S}}_{\boldsymbol{V_M}})^{-1})_{\boldsymbol{uv}} = \mathbf{0}$. This establish that $(\hat{\boldsymbol{S}}_{\boldsymbol{V_M}})^{-1} \in \mathbf{PD}^0(\boldsymbol{\mathcal{U}_M})$ and finally that $\hat{\boldsymbol{S}}_{\boldsymbol{V_M}} \in \mathbf{PD}(\boldsymbol{\mathcal{U}_M})$. Note then that $\boldsymbol{p}(\hat{\boldsymbol{S}}_{\boldsymbol{V_M}}) = \boldsymbol{S}_{\boldsymbol{V_M}}$. We will now establish that the image of $(\boldsymbol{S}_{[\boldsymbol{M}]\bullet}, \boldsymbol{S}_{[\boldsymbol{M}>\bullet}, \boldsymbol{S}_{\boldsymbol{V_M}})$ by the mapping (??) is $\boldsymbol{S}$. Only the $[\boldsymbol{M}] \times \boldsymbol{V_M}$ and the $[\boldsymbol{M}] \times [\boldsymbol{M}]$ sub-matrices of the image has to be checked. Define $\boldsymbol{R} := \boldsymbol{V_M} \setminus < \boldsymbol{M} >$. Then $\boldsymbol{V} = [\boldsymbol{M}] \dot\cup < \boldsymbol{M} > \dot\cup \boldsymbol{R}$ and $(\hat{\boldsymbol{S}}^{-1})_{[\boldsymbol{M}]\times\boldsymbol{R}} = \mathbf{0}$. Thus the $([\boldsymbol{M}]\dot\cup\boldsymbol{R}) \times ([\boldsymbol{M}]\dot\cup\boldsymbol{R})$ matrix $(\hat{\boldsymbol{S}}^{-1})_{[\boldsymbol{M}]\dot\cup\boldsymbol{R}}$ is a $\mathbf{2} \times \mathbf{2}$ block diagonal matrix. Since
$(\hat{\boldsymbol{S}}^{-1})_{[\boldsymbol{M}]\dot\cup\boldsymbol{R}} = (\hat{\boldsymbol{S}}_{[\boldsymbol{M}]\dot\cup\boldsymbol{R}} - \hat{\boldsymbol{S}}_{([\boldsymbol{M}]\dot\cup\boldsymbol{R})\times<\boldsymbol{M}>}(\hat{\boldsymbol{S}}_{<\boldsymbol{M}>})^{-1}\hat{\boldsymbol{S}}^{\boldsymbol{t}}_{([\boldsymbol{M}]\dot\cup\boldsymbol{R})\times<\boldsymbol{M}>})^{-1}$ the $([\boldsymbol{M}]\dot\cup\boldsymbol{R}) \times ([\boldsymbol{M}]\dot\cup\boldsymbol{R})$ matrix $\hat{\boldsymbol{S}}_{[\boldsymbol{M}]\dot\cup\boldsymbol{R}} - \hat{\boldsymbol{S}}_{([\boldsymbol{M}]\dot\cup\boldsymbol{R})\times<\boldsymbol{M}>}(\hat{\boldsymbol{S}}_{<\boldsymbol{M}>})^{-1}\boldsymbol{S}^{\boldsymbol{t}}_{([\boldsymbol{M}]\dot\cup\boldsymbol{R})\times<\boldsymbol{M}>}$ is also block diagonal. Thus we have $\hat{\boldsymbol{S}}_{[\boldsymbol{M}]\times\boldsymbol{R}} = \hat{\boldsymbol{S}}_{[\boldsymbol{M}>}\hat{\boldsymbol{S}}^{-1}_{<\boldsymbol{M}>}\hat{\boldsymbol{S}}_{<\boldsymbol{M}>\times\boldsymbol{R}}$. For the $[\boldsymbol{M}]\times[\boldsymbol{M}]$ sub-matrix we get $\hat{\boldsymbol{S}}_{[\boldsymbol{M}]\bullet} + \hat{\boldsymbol{S}}_{[\boldsymbol{M}>\bullet\mathbf{0}}\hat{\boldsymbol{S}}_{\boldsymbol{V_M}}\hat{\boldsymbol{S}}^{\boldsymbol{t}}_{[\boldsymbol{M}>\bullet\mathbf{0}} = \hat{\boldsymbol{S}}_{[\boldsymbol{M}]\bullet} + \hat{\boldsymbol{S}}_{[\boldsymbol{M}>\bullet}\hat{\boldsymbol{S}}_{<\boldsymbol{M}>}\hat{\boldsymbol{S}}^{\boldsymbol{t}}_{[\boldsymbol{M}>\bullet} = \hat{\boldsymbol{S}}_{[\boldsymbol{M}]}$. For the $[\boldsymbol{M}] \times \boldsymbol{V_M}$ sub-matrix we get $(\hat{\boldsymbol{S}}_{[\boldsymbol{M}>\bullet\mathbf{0}}\hat{\boldsymbol{S}}_{\boldsymbol{V_M}})_{[\boldsymbol{M}]\times<\boldsymbol{M}>} = \hat{\boldsymbol{S}}_{[\boldsymbol{M}>\bullet}\hat{\boldsymbol{S}}_{<\boldsymbol{M}>} = \hat{\boldsymbol{S}}_{[\boldsymbol{M}>}$ and $(\hat{\boldsymbol{S}}_{[\boldsymbol{M}>\bullet\mathbf{0}}\hat{\boldsymbol{S}}_{\boldsymbol{V_M}})_{[\boldsymbol{M}]\times\boldsymbol{R}} = \hat{\boldsymbol{S}}_{[\boldsymbol{M}>}\hat{\boldsymbol{S}}^{-1}_{<\boldsymbol{M}>}\hat{\boldsymbol{S}}_{<\boldsymbol{M}>\times\boldsymbol{R}} = \hat{\boldsymbol{S}}_{[\boldsymbol{M}]\times\boldsymbol{R}}$. □

Proposition 0.3. *The Jacobian of the mapping* (**??**) *and its inverse are* $|(\boldsymbol{S_M})_{<M>}|^{[M]}$ *and* $|\boldsymbol{S}_{<M>}|^{-[M]}$, *respectively.*

Proof. Partition $\boldsymbol{S} \equiv \boldsymbol{S}(\boldsymbol{\Lambda_M}, \boldsymbol{R}_{[M>}, \boldsymbol{S_M}) \in \mathbf{P}(\boldsymbol{\mathcal{U}})$ into the three ordered components $(\boldsymbol{S}_{\mathbf{V}_M}, \boldsymbol{S}_{[M>}, \boldsymbol{S}_{[M]})$ all function of $(\boldsymbol{\Lambda_M}, \boldsymbol{R}_{[M>}, \boldsymbol{S_M}) \in \mathbf{PD}([M]) \times \mathbb{R}^{[B]\times<B>} \times \mathbf{P}(\boldsymbol{\mathcal{U}_M})$. Since $\boldsymbol{S}_{\mathbf{V}_M}(\boldsymbol{\Lambda_M}, \boldsymbol{R}_{[M>}, \boldsymbol{S_M}) = \boldsymbol{S_M}$ and $\boldsymbol{S}_{[M>}(\boldsymbol{\Lambda_M}, \boldsymbol{R}_{[M>}, \boldsymbol{S_M}) = (\boldsymbol{R}_{[M>0}\widehat{\boldsymbol{S_M}})_{[M>} = \boldsymbol{R}_{[M>}(\widehat{\boldsymbol{S_M}})_{<M>}$ we get

$$\frac{\boldsymbol{d}(\boldsymbol{S}_{\mathbf{V}_M}, \boldsymbol{S}_{[M>}, \boldsymbol{S}_{[M]})}{\boldsymbol{d}(\boldsymbol{\Lambda_M}, \boldsymbol{R}_{[M>}, \boldsymbol{S_M})} = \begin{pmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} \\ * & \mathbf{1}_{[M]} \otimes (\widehat{\boldsymbol{S_M}})_{<M>} & \mathbf{0} \\ * & * & \mathbf{1} \end{pmatrix},$$

where the 1's represent identity mappings and the *'s represent entries that are not necessary to calculate. Thus the first half of the proposition follows. For the inverse mapping we have $\boldsymbol{S_M}(\boldsymbol{S}) = \boldsymbol{S}_{\mathbf{V}_M}$. Thus $\widehat{\boldsymbol{S}}_{\boldsymbol{M}}(\boldsymbol{S}) = \widehat{\boldsymbol{S}_{\mathbf{V}_M}} = (\hat{\boldsymbol{S}})_{\mathbf{V}_M}$ we have that $(\widehat{\boldsymbol{S}}_{\boldsymbol{M}}(\boldsymbol{S}))_{<M>} = (\hat{\boldsymbol{S}})_{<M>}$. Thus the second result follows. □

The fundamental integral.

Let $\boldsymbol{\Delta} \in \mathrm{PD}^0(\mathcal{U})$ and let $\boldsymbol{\lambda} \equiv (\boldsymbol{\lambda}_B | B \in V/\sim) \in \mathbb{R}_+^{V/\sim}$.

Consider the integral

$$J_{\mathcal{V}}(\Delta, \lambda) := \int_{\mathrm{P}(\mathcal{U})} \prod (|S_{[B]\bullet}|^{\lambda_B} | B \in V/\sim) \exp\{-\mathrm{Tr}(\Delta S)\} d\nu_{\mathcal{V}}(S),$$

where

$$d\nu_{\mathcal{V}}(S) := \prod (|S_{[B]\bullet}|^{-\frac{[B]+<B>+1}{2}} |S_{<B>}|^{-\frac{[B]}{2}} | B \in V/\sim) dS$$

The integral will now be transformed by (??) to an integral on the righthand side of (??). Note that the measure $\nu_{\mathcal{V}}$, by Proposition ??, is transformed into the measure

$$|L_M|^{-\frac{[B]+<B>+1}{2}} |(S_M)_{<M>}|^{\frac{[M]}{2}} dL_M dR_M d\nu_{\mathcal{V}_M}(S_M)$$

We now rewrite the function to be integrated in terms of the variables on the left hand side of (??):

Let

$$\boldsymbol{S} = \boldsymbol{p}(\begin{pmatrix} \mathbf{1}_{V_M} & \mathbf{0} \\ \boldsymbol{R}_{M0} & \mathbf{1}_M \end{pmatrix} \begin{pmatrix} \hat{\boldsymbol{S}}_M & \mathbf{0} \\ \mathbf{0} & \boldsymbol{L}_M \end{pmatrix} \begin{pmatrix} \mathbf{1}_{V_M} & \boldsymbol{R}^t_{M0} \\ \mathbf{0} & \mathbf{1}_M \end{pmatrix}) \tag{0.8}$$

and

$$\boldsymbol{\Delta} = \begin{pmatrix} \mathbf{1}_{V_M} & \boldsymbol{P}^t_{M0} \\ \mathbf{0} & \mathbf{1}_M \end{pmatrix} \begin{pmatrix} \boldsymbol{\Delta}_M & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Upsilon}_M \end{pmatrix} \begin{pmatrix} \mathbf{1}_{V_M} & \mathbf{0} \\ \boldsymbol{P}_{M0} & \mathbf{1}_M \end{pmatrix} \tag{0.9}$$

The fundamental integral, cont.

Then

$$\mathrm{Tr}(\Delta S)) = \mathrm{Tr}(\Delta \hat{S})) =$$

$$\mathrm{Tr}(\begin{pmatrix} 1_{V_M} & P^t_{M0} \\ 0 & 1_M \end{pmatrix} \begin{pmatrix} (\Delta_M & 0 \\ 0 & \Upsilon_M \end{pmatrix} \begin{pmatrix} 1_{V_M} & 0 \\ P_{M0} & 1_M \end{pmatrix} \cdot$$

$$\begin{pmatrix} 1_{V_M} & 0 \\ R_{M0} & 1_{[M]} \end{pmatrix} \begin{pmatrix} \hat{S}_M & 0 \\ 0 & L_M \end{pmatrix} \begin{pmatrix} 1_{V_M} & R^t_{M0} \\ 0 & 1_M \end{pmatrix}) =$$

$$\mathrm{Tr}(\begin{pmatrix} \Delta_M & 0 \\ 0 & \Upsilon_M \end{pmatrix} \cdot$$

$$\begin{pmatrix} 1_{V_M} & 0 \\ R_{M0} + P_{M0} & 1_M \end{pmatrix} \begin{pmatrix} \hat{S}_M & 0 \\ 0 & L_M \end{pmatrix} \begin{pmatrix} 1_{V_M} & R^t_{M0} + P^t_{M0} \\ 0 & 1_M \end{pmatrix}) =$$

$$\mathrm{Tr}(\begin{pmatrix} \Delta_M & 0 \\ 0 & \Upsilon_M \end{pmatrix} \begin{pmatrix} \hat{S}_M & 0 \\ 0 & L_M + (R_{M0} + P_{M0})\hat{S}_M(R_{M0} + P_{M0})^t \end{pmatrix}) =$$

$$\mathrm{Tr}(\Delta_M S_M) + \mathrm{Tr}(\Upsilon_M L_M) + \mathrm{Tr}(\Upsilon_M (R_{M0} + P_{M0})\hat{S}_M (R_{M0} + P_{M0})^t).$$

The transformed integral is therefore

$$\int_{\mathbf{PD}([M])\times\mathbb{R}^{[M]\times<M>}\times\mathbf{P}(\mathcal{U}_M)} |\boldsymbol{L}_M|^{\boldsymbol{\lambda}_M - \frac{[M]+<M>+1}{2}} \exp\{-\mathrm{Tr}(\boldsymbol{\Upsilon}_M \boldsymbol{L}_M)\}\cdot$$

$$|(\boldsymbol{S}_M)_{<M>}|^{\frac{[M]}{2}} \exp\{-\mathrm{Tr}(\boldsymbol{\Upsilon}_M(\boldsymbol{R}_M - \boldsymbol{P}_M)(\hat{\boldsymbol{S}}_M)_{<M>}(\boldsymbol{R}_M - \boldsymbol{P}_M^t))\}\cdot$$

$$\prod(|(\hat{\boldsymbol{S}}_M)_{[B]\bullet}|^{\lambda_B}|B \in \boldsymbol{V}_M/\sim) \exp\{-\mathrm{Tr}(\boldsymbol{\Delta}_M \boldsymbol{S}_M)\} d\boldsymbol{L}_M d\boldsymbol{R}_M d\boldsymbol{\nu}_{\boldsymbol{\mathcal{V}}_M}(\boldsymbol{S}_M) =$$

$$\pi^{\frac{[M]([M]-1)}{4}} \prod(\Gamma(\boldsymbol{\lambda}_M - \frac{<M>}{2} - \frac{i-1}{2})|i = 1, \cdots, [M])\cdot$$

$$|\boldsymbol{\Upsilon}_M|^{-\boldsymbol{\lambda}_M + \frac{<M>}{2}} \pi^{\frac{[M]<M>}{2}} |\boldsymbol{\Upsilon}_M|^{-\frac{<M>}{2}} \boldsymbol{J}_{\boldsymbol{\mathcal{V}}_M}(\boldsymbol{\Delta}_M, \boldsymbol{\lambda}_M) =$$

$$|\boldsymbol{\Upsilon}_M|^{-\boldsymbol{\lambda}_M} \boldsymbol{J}_{\boldsymbol{\mathcal{V}}_M}(\boldsymbol{\Delta}_M, \boldsymbol{\lambda}_{-M}),$$

where $\boldsymbol{\lambda}_{-M} = (\boldsymbol{\lambda}_B | \boldsymbol{B} \in \boldsymbol{V}_M/\sim)$.

Continuing descending along a never decreasing ordering of elements in $V/\sim$ and noting that $\Upsilon_M = \Delta_{M^\circ}$ we obtain that the integral converges if and only if

$$\lambda_B > \frac{[B] + < B > -1}{2},\ B \in V/\sim \tag{0.10}$$

and the value is

$$J_{\mathcal{V}}(\Delta, \lambda) = c_{\mathcal{V}}(\lambda) \prod(|\Delta_{[B]^\circ}|^{-\lambda_B} | B \in V/\sim),$$

where

$$c_{\mathcal{V}}(\lambda) := \pi^{\frac{\mathrm{Dim}(\mathrm{P}(\mathcal{U}))-V}{2}} \prod(\prod(\Gamma(\lambda_B - \frac{< B >}{2} - \frac{i-1}{2}) | i = 1, \cdots, [B]) | B \in V/\sim).$$

Let $\boldsymbol{\lambda} \equiv (\boldsymbol{\lambda}_B | \boldsymbol{B} \in \boldsymbol{V}/\sim)$ satisfy (??). The full natural and canonical exponential family on $\mathbf{P}(\boldsymbol{\mathcal{U}}) \subset \mathrm{S}^0(\boldsymbol{\mathcal{U}})$ generated by the measure $\prod(|\hat{\boldsymbol{S}}_{[\boldsymbol{B}]\bullet}|^{\boldsymbol{\lambda}_B} | \boldsymbol{B} \in \boldsymbol{V}/\sim) \boldsymbol{d\nu_{\mathcal{V}}}(\boldsymbol{S})$ is thus

$$(\boldsymbol{R}_{\boldsymbol{\Delta},\boldsymbol{\lambda}} \in \boldsymbol{\mathcal{P}}(\mathbf{S}(\boldsymbol{\mathcal{U}})) | \boldsymbol{\Delta} \in \mathbf{PD}^0(\boldsymbol{\mathcal{U}})),$$

where $\boldsymbol{dR}_{\boldsymbol{\Delta},\boldsymbol{\lambda}}(\boldsymbol{S}) :=$

$$\frac{\boldsymbol{\pi}^{\frac{\mathrm{Dim}(\mathbf{P}(\boldsymbol{V}))-\boldsymbol{V})}{2}} \prod(|\boldsymbol{\Delta}_{[\boldsymbol{B}]\circ}|^{\boldsymbol{\lambda}_B} | \boldsymbol{B} \in \boldsymbol{V}/\sim) \prod(|\boldsymbol{S}_{[\boldsymbol{B}]\bullet}|^{\boldsymbol{\lambda}_B} | \boldsymbol{B} \in \boldsymbol{V}/\sim)}{\prod(\prod(\boldsymbol{\Gamma}(\boldsymbol{\lambda}_B - \frac{<\boldsymbol{B}>}{2} - \frac{\boldsymbol{i}-1}{2}) | \boldsymbol{i} = 1, \cdots, [\boldsymbol{B}]) | \boldsymbol{B} \in \boldsymbol{V}/\sim)} \exp\{-\mathbf{Tr}(\boldsymbol{\Delta S}) \boldsymbol{d\nu_{\mathcal{V}}}(\boldsymbol{S})$$

Note that these families are dependent of the representation of the UDG $\boldsymbol{\mathcal{U}}$ as an AMG $\boldsymbol{\mathcal{V}}$ with complete boxes and without triplexes.

Definition 0.2. The probability $\boldsymbol{R}_{\boldsymbol{\Delta},\boldsymbol{\lambda}}$ is called the *generalized Rietz distribution* on $\mathbf{P}(\boldsymbol{\mathcal{U}})$ wrt. the representation $\boldsymbol{\mathcal{V}}$ of $\boldsymbol{\mathcal{U}}$ and with *shape parameter* $\boldsymbol{\lambda} \equiv (\boldsymbol{\lambda}_B | \boldsymbol{b} \in \boldsymbol{V}/\sim)$ and *natural parameter* $\boldsymbol{\Delta}$.

Expectation of generalized Rietz distribution.

Proposition 0.4. *The expectation* $\mathbb{E}(\boldsymbol{R}_{\boldsymbol{\Delta},\boldsymbol{\lambda}})$ *of* $\boldsymbol{R}_{\boldsymbol{\Delta},\boldsymbol{\lambda}}$ *is given by*

$$\mathbb{E}(\boldsymbol{R}_{\boldsymbol{\Delta},\boldsymbol{\lambda}}) = \boldsymbol{\Delta}^{\boldsymbol{\lambda}-}.$$

Proof. We shall use induction after $\boldsymbol{V}/\sim$. Using the representations of $\boldsymbol{S}$ and $\boldsymbol{\Delta}$ given in (**??**) and (**??**), respectively, and the integral transformation given by (**??**), we get

$$\mathbb{E}(\boldsymbol{R}_{\boldsymbol{\Delta},\boldsymbol{\lambda}}) = \int_{\mathrm{P}(\mathcal{U})} \boldsymbol{S} d\boldsymbol{R}_{\boldsymbol{\Delta},\boldsymbol{\lambda}}(\boldsymbol{S}) =$$

$$\boldsymbol{p}\Big(\frac{1}{J_{\mathcal{V}}(\boldsymbol{\Delta},\boldsymbol{\lambda})}\int_{\mathrm{PD}([M])\times\mathbb{R}^{[M]\times<M>}\times\mathrm{P}(\mathcal{U}_M)}\begin{pmatrix}\hat{\boldsymbol{S}}_M & \hat{\boldsymbol{S}}_M\boldsymbol{R}_{M0}^t\\ \boldsymbol{R}_{M0}\hat{\boldsymbol{S}}_M & \boldsymbol{L}_M+\boldsymbol{R}_M(\hat{\boldsymbol{S}}_M)_{<M>}\boldsymbol{R}_M\end{pmatrix}\cdot$$

$$|\boldsymbol{L}_M|^{\boldsymbol{\lambda}_M-\frac{[M]+<M>+1}{2}}\exp\{-\mathrm{Tr}(\boldsymbol{\Upsilon}_M\boldsymbol{L}_M)\}\cdot$$

$$|(\boldsymbol{S}_M)_{<M>}|^{\frac{[M]}{2}}\exp\{-\mathrm{Tr}(\boldsymbol{\Upsilon}_M(\boldsymbol{R}_M-\boldsymbol{P}_M)(\boldsymbol{S}_M)_{<M>}(\boldsymbol{R}_M-\boldsymbol{P}_M)^t)\}\cdot$$

$$\prod(|(\boldsymbol{S}_M)_{[B]\bullet}|^{\boldsymbol{\lambda}_B}|B\in \boldsymbol{V}_M/\sim)\exp\{-\mathrm{Tr}(\boldsymbol{\Delta}_M\boldsymbol{S}_M)\}d\boldsymbol{L}_M d\boldsymbol{R}_M d\boldsymbol{\nu}_{\mathcal{V}_M}(\boldsymbol{S}_M)\Big) =$$

$$\boldsymbol{p}\Big(\begin{pmatrix}\hat{\boldsymbol{\Delta}}_M^{\boldsymbol{\lambda}_{-M-}} & \hat{\boldsymbol{\Delta}}_M^{\boldsymbol{\lambda}_{-M-}}\boldsymbol{P}_{M0}^t\\ \boldsymbol{P}_{M0}\hat{\boldsymbol{\Delta}}_M^{\boldsymbol{\lambda}_{-M-}} & \mathrm{LRC},\end{pmatrix}\Big),$$

where the lower right corner LRC is calculated by first integrating out over $\boldsymbol{L}_M$ to obtain $(\boldsymbol{\lambda}_M-\frac{<M>}{2})\boldsymbol{\Upsilon}^{-1}+\boldsymbol{R}_M(\boldsymbol{S}_M)_{<M>}\boldsymbol{R}_M^t$, then integrating over $\boldsymbol{R}_M$ to obtain $(\boldsymbol{\lambda}_M-\frac{<M>}{2})\boldsymbol{\Upsilon}^{-1}+\frac{<M>}{2}\boldsymbol{\Upsilon}^{-1}+\boldsymbol{P}_M(\boldsymbol{S}_M)_{<M>}\boldsymbol{P}_M^t$, and finally over $\boldsymbol{S}_M$ to obtain the LRC as $\boldsymbol{\lambda}_M\boldsymbol{\Upsilon}^{-1}+\boldsymbol{P}_M(\hat{\boldsymbol{\Delta}}^{\boldsymbol{\lambda}_{-M-}})_{<M>}\boldsymbol{P}_M^t$.

Proof of Expectation of generalized Rietz distribution, cont.

All together the expectation becomes

$$\boldsymbol{p}(\begin{pmatrix} \mathbf{1}_{V_M} & \mathbf{0} \\ \boldsymbol{P}_{M\mathbf{0}} & \mathbf{1}_M \end{pmatrix} \begin{pmatrix} ((\boldsymbol{\Delta}_M)^{\boldsymbol{\lambda}_{-M^-}} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\lambda}_M \boldsymbol{\Upsilon}_M^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{1}_{V_M} & \boldsymbol{P}_{M\mathbf{0}}^t \\ \mathbf{0} & \mathbf{1}_M \end{pmatrix}) = \boldsymbol{\Delta}^{\boldsymbol{\lambda}-}.$$

□

If $\boldsymbol{\Delta} \in \mathbf{PD}^0(\boldsymbol{\mathcal{U}})$ is replaced by the $\boldsymbol{\Sigma}^{-\boldsymbol{\lambda}}$, $\boldsymbol{\Sigma} \in \mathbf{P}(\boldsymbol{\mathcal{U}})$ we obtain the Rietz distributions parameterized by their expectation. Using that $(\boldsymbol{\Sigma}^{-\boldsymbol{\lambda}})_{[B]^\circ} = \boldsymbol{\lambda}_B(\boldsymbol{\Sigma}_{[B]\bullet})^{-1}$, $\boldsymbol{B} \in \boldsymbol{V}/\sim$ we obtain

$$d\mathbb{R}_{\Sigma,\lambda}(S) := \frac{\pi^{\frac{\mathrm{Dim}(\mathrm{P}(V))-V)}{2}} \prod(\lambda_B^{\lambda_B[B]} | B \in V/\sim)}{\prod(\prod(\Gamma(\lambda_B - \frac{<B>}{2} - \frac{i-1}{2}) | i = 1, \cdots, [B]) | B \in V/\sim)}.$$

$$\frac{\prod(|S_{B\bullet}|^{\lambda_B} | B \in V/\sim)}{\prod(|\Sigma_{B\bullet}|^{\lambda_B} | B \in V/\sim)} \exp\{-\mathrm{Tr}(\Sigma^{-\lambda} S)\} d\nu_G(S),$$

where $\boldsymbol{d\nu_G(S) := c \cdot \prod(|S_{B\bullet}|^{\lambda_B} | B \in V/\sim) d\nu_{\mathcal{V}}(S)}$

Definition 0.3. The probability $\mathbb{R}_{\boldsymbol{\Sigma},\boldsymbol{\lambda}}$ is called the *generalized Rietz distribution* on $\mathbf{P}(\boldsymbol{\mathcal{U}})$ wrt. the representation $\boldsymbol{\mathcal{V}}$ of $\boldsymbol{\mathcal{U}}$ and with *shape parameter* $\boldsymbol{\lambda} \equiv (\boldsymbol{\lambda}_B | \boldsymbol{B} \in \boldsymbol{V}/\sim)$ and *expectation parameter* $\boldsymbol{\Sigma}$.

The generalized Rietz model.

The *Rietz model* wrt to $\boldsymbol{\mathcal{V}}$ (with fixed shape parameter $\boldsymbol{\lambda}$) in its expectation parametrization is then

$$(\mathbb{R}_{\boldsymbol{\Sigma},\boldsymbol{\lambda}} \in \boldsymbol{\mathcal{P}}(\mathrm{S}(\boldsymbol{\mathcal{U}}))|\boldsymbol{\Sigma} \in \mathbf{P}(\boldsymbol{\mathcal{U}}))$$

It is trivial that the ML estimator $\hat{\boldsymbol{\Sigma}}(\boldsymbol{S})$ for $\boldsymbol{\Sigma} \in \mathbf{P}(\boldsymbol{\mathcal{U}})$ at the observation point $\boldsymbol{S} \in \mathbf{P}(\boldsymbol{\mathcal{U}})$ exists for all $\boldsymbol{S} \in \mathbf{P}(\boldsymbol{\mathcal{U}})$ and it is uniquely given by $\hat{\boldsymbol{\Sigma}}(\boldsymbol{S}) = \boldsymbol{S}$.

It is trivial that $\hat{\boldsymbol{\Sigma}}$ is complete and sufficient.

OK.

Connection to the work by Letac and Massam I

In this remark we investigate the relation to the work by Letac and Massam (200?) abbreviated LM. Let $\mathcal{U}$ be an DUG. In the paper LM, called *Wishart distributions for decomposable graphs*, the authors define their *Wishart distributions of type I* on the open convex cone $Q_{\mathcal{U}} = \mathrm{P}(\mathcal{U})$, to connect our notation to their notation. We now recall, in fact almost quoting, as good as possible, Letac and Massam's definitions surrounding their Wishart distributions of type I, only. Let $\mathcal{C}$, $\mathcal{S}$, and $\nu(S)$ denote the set of cliques, the set of separators, and the multiplicity of the sparator $S \in \mathcal{S}$, respectively. As in LM we shall now assume $\mathcal{C} > 1$.

The central idea in LM is the integral

$$\boldsymbol{I}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\Delta}) := \int_{\mathrm{P}(\boldsymbol{\mathcal{U}})} \boldsymbol{H}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{X}) \exp\{-\mathrm{Tr}(\boldsymbol{\Delta}\boldsymbol{X})\} d\boldsymbol{\mu}_{\mathcal{U}}(\boldsymbol{X}),$$

$(\boldsymbol{\alpha},\boldsymbol{\beta}) \equiv ((\boldsymbol{\alpha}_C|\boldsymbol{C} \in \boldsymbol{\mathcal{C}}), (\boldsymbol{\beta}_S|\boldsymbol{S} \in \boldsymbol{\mathcal{S}})) \in \mathbb{R}^{\mathcal{C}} \times \mathbb{R}^{\mathcal{S}}$, $\boldsymbol{\Delta} \in \mathrm{PD}^0(\boldsymbol{\mathcal{U}})$, where

$$\boldsymbol{H}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{X}) := \frac{\prod(|\boldsymbol{X}_C|^{\alpha_C}|\boldsymbol{C} \in \boldsymbol{\mathcal{C}})}{\prod(|\boldsymbol{X}_S|^{\nu(S)\beta_S}|\boldsymbol{S} \in \boldsymbol{\mathcal{S}})}$$

$$d\boldsymbol{\mu}_{\mathcal{U}}(\boldsymbol{X}) := \boldsymbol{H}((-\frac{\boldsymbol{C}+1}{2}|\boldsymbol{C} \in \boldsymbol{\mathcal{C}}), (-\frac{\boldsymbol{S}+1}{2}|\boldsymbol{S} \in \boldsymbol{\mathcal{S}}), \boldsymbol{X}) d\boldsymbol{X}$$

The set of $(\boldsymbol{\alpha},\boldsymbol{\beta}) \in \mathbb{R}^{\mathcal{C}} \times \mathbb{R}^{\mathcal{S}}$ such that $\boldsymbol{I}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\Delta}) < \infty$ and such that $\boldsymbol{I}(\boldsymbol{\alpha},\boldsymbol{\beta},\hat{\boldsymbol{X}}^{-1})/\boldsymbol{H}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{X}) = \boldsymbol{c}(\boldsymbol{\alpha},\boldsymbol{\beta})$, $\boldsymbol{X} \in \mathrm{P}(\boldsymbol{\mathcal{U}})$ is denoted $\boldsymbol{\mathcal{A}}$. The *Wishart distribution of type I* with parameters $((\boldsymbol{\alpha},\boldsymbol{\beta}),\boldsymbol{\Sigma}) \in \boldsymbol{\mathcal{A}} \times \mathrm{P}(\boldsymbol{\mathcal{U}})$ is then by LM defined to be

$$d\boldsymbol{W}_{\mathrm{P}(\mathcal{U}),\alpha,\beta,\Sigma}(\boldsymbol{X}) = \frac{1}{\boldsymbol{c}(\boldsymbol{\alpha},\boldsymbol{\beta})\boldsymbol{H}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\Sigma})} \boldsymbol{H}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{X}) \exp\{-\mathrm{Tr}(\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X})\} d\boldsymbol{X}.$$

Connection to the work by Letac and Massam III

The problem of characterizing $\mathcal{A}$ and calculating $c(\alpha, \beta)$, $(\alpha, \beta) \in \mathcal{A}$ is therefore main consideration in LM. Nevertheless LM do not obtain a complete solution to this problem.

If $\mathcal{U}$ does not contain the DUG $\bullet - \bullet - \bullet - \bullet$, denoted A_4 in LM, as an induced subgraph the open convex cone $\mathbf{P}(\mathcal{U})$ is a homogeneous cone. In this case the family Wishart distributions of type I is identical[‡] with the family of *general Wishart distributions* on $\mathbf{P}(\mathcal{U})$ obtained by Andersson and Wojnar (200?) for ANY homogeneous cone, in particular of course for the homogeneous cone $\mathbf{P}(\mathcal{U})$. In this case the problem is thus already solved completely. Nevertheless, LM presents a self contained version of the solution in this special case.

[‡] except for a trivial re-parametrization

Connection to the work by Letac and Massam IV

In the non-homogeneous case, i.e., $\mathcal{U}$ does contain A_4 as an induced subgraph, LM's solution to the above problem is not complete, as they also point out themselfe. They find subsets $\mathcal{A}_P \subseteq \mathcal{A}$, each subset in general depends on a perfect ordering P of the cliques, $C_1, \cdots, C_{\mathcal{C}}$ and it only has dimension $\mathcal{C}+1$, cf. below. Thus $\cup_P \mathcal{A}_P \subseteq \mathcal{A}$. Since it is unknown whether equality (probably not in general) holds $\mathcal{A}$ is therefore not characterized[§] The family of Wishart distributions of type I parameterized by $\mathcal{A}_P \times \mathrm{P}(\mathcal{U})$ is closed under convolution but depends of an arbitrary choice of perfect ordering P. If $\mathcal{A}_P$ is replaced by $\cup_P \mathcal{A}_P$ the family does not depend of P but is **not** in general closed under convolution.

Furthermore the method does not work in the homogeneous case.

We shall now for a fixed P establish that the family of Wishart distributions of type I parameterized by $\mathcal{A}_P \times \mathrm{P}(\mathcal{U})$ defined by LM is a special case of our families of Rietz distributions.

[§]Nevertheless, it is interesting that in the case of A_4 LM establish that equality holds.

Let for $S_2, S_3, \cdots, S_{\mathcal{C}}$ be the to P corresponding ordered listing of separators $S_2, S_3, \cdots, S_{\mathcal{C}}$ with possible repeats.

The complicated definition of $\mathcal{A}_P$, cf. LM, Section 3.4, is the set of all $(\alpha, \beta) \equiv ((\alpha_C | C \in \mathcal{C}), (\beta_j | S \in \mathcal{S})) \in \mathbb{R}^{\mathcal{C}} \times \mathbb{R}^{\mathcal{S}}$ such that

$$
\begin{aligned}
\alpha_C &> \tfrac{C-1}{2}, \quad C \in \mathcal{C} \\
\alpha_{C_1} + \delta_2 &> \tfrac{S_2-1}{2} \\
\sum(\alpha_{C_q} | q \in J_P(S)) - \nu(S)\beta_S &= 0, \quad S \in \mathcal{S}, \ S \neq S_2,
\end{aligned}
\tag{0.11}
$$

where $J_P(S) = \{j \in \{1, \cdots, \mathcal{C}\} | S_j = S\}$ and $\delta_2 := \sum(\alpha_{C_q} | q \in J_P(S_2)) - \nu(S_2)\beta_2$. Furthermore, still quoting LM, for $(\alpha, \beta) \in \mathcal{A}_P$

$$
c(\alpha, \beta) := \Gamma_{S_2}(\alpha_1 + \delta_2) \frac{\Gamma_{C_1}(\alpha_1)}{\Gamma_{S_2}(\alpha_1)} \prod \left(\frac{\Gamma_{C_q}(\alpha_q)}{\Gamma_{S_q}(\alpha_q)} | q = 2, \cdots, \mathcal{C} \right)
$$

with

$$
\Gamma_r(p) := \pi^{\frac{1}{4}r(r-1)} \prod \left(\Gamma(p - \frac{j-1}{2}) | j = 1, \cdots, r\right),
$$

$p \in \mathbb{R}_+$, $r \in \mathbb{N} := \{1, 2, \cdots\}$. $p > \frac{r-1}{2}$,

Connection to the work by Letac and Massam VI

Since we have perfect order P of the cliques $C_1, \cdots, C_{\mathcal{C}}$ we have: the *history* $H_j := C_1 \cup C_2 \cup \cdots, \cup C_j$, $j = 1, 2, 3, \cdots, \mathcal{C}$, the *separators* $S_j := C_j \cap H_{j-1}$, $j = 2, 3, \cdots, \mathcal{C}$, and the *remainders* $R_j = C_j \backslash H_{j-1}$, $j = 2, 3, \cdots, \mathcal{C}$. We now replacing all undirected edges between a $v \in S_j$ and a $v' \in R_j$ with an arrow $v \longrightarrow v'$. Also we replace any undirected edge between $v \in S_2$ and $v' \in C_1 \backslash S_2$ by the arrow $v \longrightarrow v'$. It is not difficult to see using the definition of the perfect order P that this assignment is without conflicts, and that the resulting MG $\mathcal{V}_P$ is acyclic without triplexes and with the $\mathcal{C}+1$ complete boxes $S_2, C_1 \backslash S_2, R_2, \cdots, R_{\mathcal{C}}$. Furthermore by definition $\mathrm{pa}(S_2) = \emptyset$, $\mathrm{pa}(C_1 \backslash S_2) = S_2$, and $\mathrm{pa}(R_j) = S_j$, $j = 2, \cdots, \mathcal{C}$.

Note also that $\mathcal{S} = \{S_j | j = 2, \cdots, \mathcal{S}\}$.

Thus with $\mathcal{V} = \mathcal{V}_P$, and with S replaced by X our

$$d\nu_{\mathcal{V}}(X) := \prod(|\hat{X}_{[B]\bullet}|^{-\frac{[B]+<B>+1}{2}}|\hat{X}_{<B>}|^{-\frac{[B]}{2}}|B \in V/\sim)dX =$$
$$|X_{S_2}|^{-\frac{S_2+1}{2}}|X_{(C_1\setminus S_2)\bullet}|^{-\frac{C_1\setminus S_2+S_2+1}{2}}|\hat{X}_{S_2}|^{-\frac{C_1\setminus S_2}{2}}\cdot$$
$$|X_{S_2}|^{-\frac{C_1+1}{2}}|X_{(C_1\setminus S_2)\bullet}|^{-\frac{C_1+1}{2}}\prod(|X_{R_j\bullet}|^{-\frac{R_j+S_j+1}{2}}|\hat{X}_{S_j}|^{-\frac{R_j}{2}}|j = 2,\cdots,\mathcal{C})dX =$$
$$|X_{C_1}|^{-\frac{C_1+1}{2}}\prod(\left(|X_{R_j\bullet}||\hat{X}_{S_j}|\right)^{-\frac{R_j+S_j+1}{2}}|\hat{X}_{S_j}|^{+\frac{S_j+1}{2}}|j = 2,\cdots,\mathcal{C})dX =$$
$$\frac{\prod(|\hat{X}_{C_j}|^{-\frac{C_j+1}{2}}|j = 1,\cdots,\mathcal{C})}{\prod(|\hat{X}_{S_j}|^{-\frac{S_j+1}{2}}|j = 2,\cdots,\mathcal{C})}dX,$$

the measure $\mu_{\mathcal{U}}$ defined by LM.

Connection to the work by Letac and Massam VIII

Next let $(\boldsymbol{\alpha},\boldsymbol{\beta}) \equiv ((\boldsymbol{\alpha}_C|\boldsymbol{C}\in\boldsymbol{\mathcal{C}}),(\boldsymbol{\beta}_S|\boldsymbol{S}\in\boldsymbol{\mathcal{S}})) \in \mathbb{R}^{\mathcal{C}}\times\mathbb{R}^{\mathcal{S}}$ satisfying (??). Then

$$\boldsymbol{H}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{X}) = \frac{\prod(|\boldsymbol{X}_C|^{\alpha_C}|\boldsymbol{C}\in\boldsymbol{\mathcal{C}})}{\prod(|\boldsymbol{X}_S|^{\nu(S)\beta_S}|\boldsymbol{S}\in\boldsymbol{\mathcal{S}})} = \frac{\prod(|\boldsymbol{X}_{C_j}|^{\alpha_{C_j}}|\boldsymbol{j}=\boldsymbol{1},\cdots,\boldsymbol{\mathcal{C}})}{\prod(|\boldsymbol{X}_{S_j}|^{\beta_{S_j}}|\boldsymbol{j}=\boldsymbol{2},\cdots,\boldsymbol{\mathcal{C}})} =$$
$$|\boldsymbol{X}_{C_1}|^{\alpha_1}\prod(|\boldsymbol{X}_{R_j\bullet}|^{\alpha_{C_j}}|\boldsymbol{j}=\boldsymbol{2},\cdots,\boldsymbol{\mathcal{C}})\prod(|\boldsymbol{X}_{S_j}|^{\alpha_{C_j}-\beta_{S_j}}|\boldsymbol{j}=\boldsymbol{2},\cdots,\boldsymbol{\mathcal{C}}) =$$
$$|\boldsymbol{X}_{C_1}|^{\alpha_1}\prod(|\boldsymbol{X}_{R_j\bullet}|^{\alpha_{C_j}}|\boldsymbol{j}=\boldsymbol{2},\cdots,\boldsymbol{\mathcal{C}})\prod(\prod(|\boldsymbol{X}_S|^{\alpha_{C_j}-\beta_S}|\boldsymbol{j}\in\boldsymbol{J_P}(\boldsymbol{S}))|\boldsymbol{S}\in\boldsymbol{\mathcal{S}}) =$$
$$|\boldsymbol{X}_{C_1}|^{\alpha_1}|\boldsymbol{X}_{S_2}|^{\delta_2}\prod(|\boldsymbol{X}_{R_j\bullet}|^{\alpha_{C_j}}|\boldsymbol{j}=\boldsymbol{2},\cdots,\boldsymbol{\mathcal{C}})\prod(|\boldsymbol{X}_S|^{\sum(\alpha_{C_j}|j\in J_P(S))-\nu(S)\beta_S}|\boldsymbol{S}\in\boldsymbol{\mathcal{S}}\backslash\boldsymbol{S}_2)$$
$$|\boldsymbol{X}_{C_1\backslash S_2\bullet}|^{\alpha_1}|\boldsymbol{X}_{S_2}|^{\alpha_1+\delta_2}\prod(|\boldsymbol{X}_{R_j\bullet}|^{\alpha_{C_j}}|\boldsymbol{j}=\boldsymbol{2},\cdots,\boldsymbol{\mathcal{C}})\cdot\boldsymbol{1} =$$
$$\prod(|\boldsymbol{X}_{B\bullet}|^{\lambda_B}\boldsymbol{B}\in\boldsymbol{V}/\sim)$$

with

$$\boldsymbol{\lambda}_B = \begin{cases} \boldsymbol{\alpha}_1+\boldsymbol{\delta}_2 & \text{for} \quad \boldsymbol{B}=\boldsymbol{S}_2 \\ \boldsymbol{\alpha}_{C_1} & \text{for} \quad \boldsymbol{B}=\boldsymbol{C}_1\backslash\boldsymbol{S}_2 \\ \boldsymbol{\alpha}_{C_j} & \text{for} \quad \boldsymbol{B}=\boldsymbol{R}_j,\ \boldsymbol{j}=\boldsymbol{2},\cdots,\boldsymbol{\mathcal{C}} \end{cases}, \tag{0.12}$$

an apparently much simpler expression.

Furthermore we have

$$\frac{B+<B>-1}{2} = \begin{cases} \frac{S_2+\emptyset-1}{2} & \text{for} \quad B = S_2 \\ \frac{C_1\backslash S_2+S_2-1}{2} & \text{for} \quad B = C_1\backslash S_2 \\ \frac{R_j+S_j-1}{2} & \text{for} \quad B = R_j,\ j = 2,\cdots,\mathcal{C} \end{cases} =$$

$$\begin{cases} \frac{S_2-1}{2} & \text{for} \quad B = S_2 \\ \frac{C_1-1}{2} & \text{for} \quad B = C_1\backslash S_2 \\ \frac{C_j-1}{2} & \text{for} \quad B = R_j,\ j = 2,\cdots,\mathcal{C} \end{cases}.$$

Thus the the two last conditions in (??) are equivalent with our (??).

Connection to the work by Letac and Massam X

For a fixed perfect ordering $\boldsymbol{P}$ of the cliques this established that the family Wishart distribution of type I, $\boldsymbol{W}_{\mathrm{P}(\boldsymbol{\mathcal{U}}),\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\sigma}}$, $((\boldsymbol{\alpha},\boldsymbol{\beta}),\boldsymbol{\sigma}) \in \boldsymbol{\mathcal{A}}_{\boldsymbol{P}} \times \mathrm{P}(\boldsymbol{\mathcal{U}})$, is the family of Rietz distribution $\boldsymbol{R}_{\boldsymbol{\Sigma},\boldsymbol{\lambda}}$,

$$(\boldsymbol{\Sigma},\boldsymbol{\lambda}) \in \mathrm{P}(\boldsymbol{\mathcal{U}}) \times (\times(]\frac{[\boldsymbol{B}] + < \boldsymbol{B} > -\mathbf{1}}{\mathbf{2}}, \infty[|\boldsymbol{B} \in \boldsymbol{V}/\sim))$$

wrt to $\boldsymbol{\mathcal{V}}_{\boldsymbol{P}}$. The one to one correspondence between the two index sets $\boldsymbol{\mathcal{A}}_{\boldsymbol{P}} \times \mathrm{P}(\boldsymbol{\mathcal{U}})$ and the above is given by (??) and $\boldsymbol{\Sigma} = (\boldsymbol{\sigma}^{-1})^{-\boldsymbol{\lambda}}$. In particular $\boldsymbol{\mathcal{A}}_{\boldsymbol{P}}$ has dimension $\boldsymbol{V}/\sim = \boldsymbol{\mathcal{C}} + \mathbf{1}$. Note that our Rietz distributions has $\mathbb{E}(\boldsymbol{R}_{\boldsymbol{\Sigma},\boldsymbol{\lambda}}) = \boldsymbol{\Sigma}$ while for the Wishart distributions of type I $\mathbb{E}(\boldsymbol{W}((\boldsymbol{\alpha},\boldsymbol{\beta}),\boldsymbol{\sigma}))$ in general is different from $\boldsymbol{\sigma}$. From our point of view only the natural parametrization or the parametrization by the expectation should be used.

Also it seems a little strange to use the functions $\boldsymbol{H}((\boldsymbol{\alpha},\boldsymbol{\beta}),\cdot)$, indexed by $(\boldsymbol{\alpha},\boldsymbol{\beta}) \in \mathbb{R}^{\mathcal{C}+\mathcal{S}}$, a $\mathcal{C}+\mathcal{S}$ dimensional vector space, when the actual set(s) $\mathcal{A}_P$ of interest are of dimension $\mathcal{C}+1$. I suggest that one takes the ordering serious and start with a representation of $\mathcal{U}$ as a mixed graph $\mathcal{V}$ with complete boxes and with no triplexes, a representation of $\mathcal{U}$, and replaces LM's "$\boldsymbol{H}$-function" by

$$(\boldsymbol{S},\boldsymbol{\lambda}) \rightarrow \prod(|\boldsymbol{S}_{\boldsymbol{B}\bullet}|^{\lambda_B}|\boldsymbol{B} \in \boldsymbol{V}/\sim)$$

$\boldsymbol{\lambda} \equiv (\boldsymbol{\lambda}_B|\boldsymbol{B} \in \boldsymbol{V}/\sim) \in \mathbb{R}^{V/\sim}$, $\boldsymbol{S} \in \mathrm{P}(\mathcal{U})$, and use our measure $\boldsymbol{\nu}_{\mathcal{V}}$ in more general depending on the representation $\mathcal{V}$ of $\mathcal{U}$. In that way one obtains as we have seen the Rietz distributions a much general class of distributions that the Wishart distributions of type I. In fact this point of view also works unaltered in the homogeneous case. In particular, we think that the Wishart distributions of Type I in LM should, since they are a subclass of our Rietz distributions, be called Rietz distributions since they depends on a representation $\mathcal{V}_P$ of $\mathcal{U}$ induced by the ordering $\boldsymbol{P}$.

From Rietz to Wishart.

From our point of view the following question natural arises:

Do that exist an intrinsic (canonical) choice of representation of a DUG $\mathcal{U}$ as a AMG $\mathcal{V}_{\mathcal{U}}$ with complete boxes and no triplexes?

The construction of $\mathcal{V}_{\mathcal{U}}$ should not depend of any kind arbitrary choice, for example a choice perfect ordering of the cliques, as in LM. The answer is yes and it therefore tempting to call the Rietz distributions wrt $\mathcal{V}_{\mathcal{U}}$ for Wishart distribution wrt $\mathcal{U}$. We will describe the construction of $\mathcal{V}_{\mathcal{U}}$.

A natural representation of a DUG as an AMG.

The construction is based on a natural partial ordering $\prec_{\mathcal{U}} \equiv \prec$ of on V generated by the DUG $\mathcal{U}$. Define for $u, v \in V$ with $u \neq v$

$$v \prec_{\mathcal{U}} u \ \text{ if } \ \{u\} \cup \text{nb}_{\mathcal{U}}(u) \subset \{v\} \cup \text{nb}_{\mathcal{U}}(v).$$

It is obvious that $\prec$ is a partial ordering of V. Note that $u \prec v$ implies that $u - v$ in $\mathcal{U}$. Also note that the relation $\prec_{\mathcal{U}}$ is empty if and only if $\mathcal{U}$ is a disjoint union of complete UGs. In that case the construction is finished (before it ever started) and $\mathcal{V}_{\mathcal{U}} = \mathcal{U}$.

Step 1a: For all $u, v \in V$ with $v \prec u$ replace the line $v - u$ with the arrow $v \longrightarrow u$. A MG without triplexes is thus obtained.

Step 1b: If an arrow $v \longrightarrow v'$ participate in a partly directed cycle replace it back to a line $v - v'$.

The order of the replacements in Step 1b, if any, does not matter. The resulting graph $\mathcal{V}_1$ is an AMG without triplexes. Since $\mathcal{U}$ is a DUG and not a disjoint union of complete UG, some lines are in fact replaced by arrows after Step 1b.

All the boxes in $\mathcal{V}_1$ are DUG as induced subgraphs of $\mathcal{V}_1$.

A natural representation of a DUG as an AMG, Step 2.

The definition of $\prec_{\mathcal{U}}$ make sense unchanged when $\boldsymbol{\mathcal{U}}$ is replaced by $\boldsymbol{\mathcal{V}}_1$. Again $\boldsymbol{v} \prec_{\mathcal{V}_1} \boldsymbol{u}$ implies $\boldsymbol{v} -_{\mathcal{V}_1} \boldsymbol{u}$ and the relation $\prec_{\mathcal{V}_1}$ is empty if and only if all boxes $\boldsymbol{B} \in \boldsymbol{V}/\sim$ are complete. In that case the construction is finished and $\boldsymbol{\mathcal{V}}_{\mathcal{U}} = \boldsymbol{\mathcal{V}}_1$. Otherwise continue with

Step 2a: For all $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{V}$ with $\boldsymbol{v} \prec_{\mathcal{V}_1} \boldsymbol{u}$ replace the line $\boldsymbol{v} -_{\mathcal{V}_1} \boldsymbol{u}$ with the arrow $\boldsymbol{v} \longrightarrow \boldsymbol{u}$. A MG without triplexes is thus again obtained.

Step 2b: If an arrow $\boldsymbol{v} \longrightarrow \boldsymbol{v}'$ participate in a partly directed cycle replace it back to a line $\boldsymbol{v} - \boldsymbol{v}'$.

Only *new* arrows from Step 2a could participate in partly directed cycles, in particular all old arrows remains.

The order of the replacements in Step 2b, if any, does not matter.

The resulting graph $\boldsymbol{\mathcal{V}}_2$ is an AMG without triplexes.

Since all boxes in $\boldsymbol{\mathcal{V}}_1$ are DUG as induced subgraphs and not all boxes in $\boldsymbol{\mathcal{V}}_1$ are complete, some of the new arrows remains after from Step 2a remains after Step 2b.

All boxes in $\boldsymbol{\mathcal{V}}_2$ are DUG as induced subgraphs.

A natural representation of a DUG as an AMG, Step k.

If all boxes in $\mathcal{V}_2$ are complete the construction is finished and $\mathcal{V}_{\mathcal{U}} = \mathcal{V}_2$. Otherwise continue with

⋮

Step ka: For all $u, v \in V$ with $v \prec_{\mathcal{V}_{k-1}} u$ replace the line $v -_{\mathcal{V}_{k-1}} u$ with the arrow $v \longrightarrow u$. A MG without triplexes is thus again obtained.

Step kb: If an arrow $v \longrightarrow v'$ participate in a partly directed cycle replace it back to a line $v - v'$.

Only *new* arrows from Step ka could participate in partly directed cycles, in particular all old arrows remains.

The order of the replacements in Step kb, if any, does not matter.

The resulting graph $\mathcal{V}_k$ is an AMG without triplexes.

Since all boxes in $\mathcal{V}_{k-1}$ are DUG as induced subgraphs and not all boxes in $\mathcal{V}_{k-1}$ are complete, some of the new arrows remains after from Step ka remains after Step kb.

All boxes in $\mathcal{V}_k$ are DUG as induced subgraphs.

If all boxes in $\mathcal{V}_k$ are complete the construction is finished and $\mathcal{V}_{\mathcal{U}} = \mathcal{V}_k$. Otherwise continue with step $k+1$, $k = 1, 2, \cdots$.

This process will terminate after finitely many steps, say n steps, ending with an AMG $\mathcal{V}_{\mathcal{U}} := \mathcal{V}_n$ with complete boxes and no triplexes.

Definition 0.4. The Rietz distributions $\mathbb{R}_{\Sigma,\lambda}$ (or $R_{\Delta,\lambda}$) on $\mathbf{P}(\mathcal{U})$ wrt. the representation $\mathcal{V}_{\mathcal{U}}$ of $\mathcal{U}$, with shape parameter $\lambda \equiv (\lambda_B | B \in V/\sim)$, and expectation parameter $\Sigma \in \mathbf{P}(\mathcal{U})$ (or natural parameter $\Delta \in \mathbf{PD}^0(\mathcal{U})$) is also called the Wishart distribution $\mathbb{W}_{\Sigma,\lambda} := \mathbb{R}_{\Sigma,\lambda}$ ($W_{\Delta,\lambda} := \mathbb{R}_{\Delta,\lambda}$) on on $\mathbf{P}(\mathcal{U})$ with *shape parameter* λ, and *expectation parameter* Σ (or *natural parameter* Δ).

The *Wishart model* (with fixed shape parameter λ) in its expectation parametrization is then

$$(\mathbb{W}_{\Sigma,\lambda} \in \mathcal{P}(\mathbf{P}(\mathcal{U})) | \Sigma \in \mathbf{P}(\mathcal{U})).$$

Outlook.

- Testing problems within generalized Rietz/Wishart distributions.
- Inverse generalized Rietz/Wishart distributions on $\mathbf{PD}^0(\mathcal{U})$.
- Expectation of inverse generalized Rietz/Wishart distributions on $\mathbf{PD}^0(\mathcal{U})$.
- Variance of Rietz/Wishart distributions.
- Generalized Rietz/ Wishart distributions on $\mathbf{PD}^0(\mathcal{U})$.
- Generalized Inverse Rietz/ Wishart distributions on $\mathbf{P}(\mathcal{U})$.
- ETC.