Supplementary Material for the paper "MRA-based Statistical Learning from Incomplete Rankings"

Eric Sibony ERIC.SIBONY@TELECOM-PARISTECH.FR
Stéphan Clémençon STEPHAN.CLEMENCON@TELECOM-PARISTECH.FR
Jérémie Jakubowicz JEREMIE.JAKUBOWICZ@TELECOM-SUDPARIS.E

1. Illustrative example of the statistical and computational challenge

We begin with an example to illustrate the complexity of the combinatorial relationships that exist between the marginals of a ranking model, and how it leads to a statistical and computational challenge. Let $n = 4$ and $\mathcal{A} = \{\{1,3\},\{2,4\},\{3,4\},\{1,2,3\},\{1,3,4\}\}$. Assuming the P_A's are known, finding a function $q \in L(\mathfrak{S}_4)$ such that $M_A q = P_A$ for all $A \in \mathcal{A}$ boils down to solving the linear system given in Figure 1. The values of q are denoted by q_σ instead of $q(\sigma)$ for $\sigma \in \mathfrak{S}_4$. This system shows that all the equations are quite entangled. For instance, the unknown q_{1234} appears in 5 equations. Hence, not only the dimension of the system quickly explodes with n, but all the equations have complex relationships, and decomposing this system into simpler ones is far from being obvious.

In a statistical setting, a natural approach would certainly be to perform a least-square regression with the unknowns as parameters. The complex relationships between the marginals would however remain, and the computation of the gradient would quickly become intractable as n grows.

2. General definitions and results

Here we introduce some general definitions and results that are useful for the technical proofs in the sequel. We denote by $\mathbb{I}\{\mathcal{E}\}$ the indicator function of any event $\mathcal{E}$ so that $\mathbb{I}\{\mathcal{E}\} = 1$ if $\mathcal{E}$ is true and 0 if it is false.

Definition 1 (Induced ranking)**.** Let $\pi \in \Gamma_n$ be an incomplete ranking and $A \in \mathcal{P}(c(\pi))$ be a subset of items in the content of π. The ranking induced by π over A is by definition the unique subword of π of content A. We denote it by $\pi_{|A}$.

Definition 2. For a ranking $\pi = \pi_1 \dots \pi_k \in \Gamma_n$ and for $1 \le i < j \le k$, we denote by $\pi_{[\![i,j]\!]}$ its subword defined by $\pi_{[\![i,j]\!]} = \pi_i \dots \pi_j$.

Definition 3. Define the coefficients $\alpha_B(\pi,\pi') := X_B \delta_\pi(\pi')$ for $B \in \mathcal{P}([\![n]\!])$ and $\pi, \pi' \in \Gamma(B)$, so that for $F \in L(\Gamma(A))$ with $A \in \mathcal{P}([\![n]\!])$ and $\pi \in \Gamma(B)$:

$$X_B F(\pi) = \sum_{\pi' \in \Gamma(B)} \alpha_B(\pi,\pi') M_B F(\pi').$$

Lemma 1. *Let $A \in \mathcal{P}([\![n]\!])$ with $|A| = k$ and $(F_B)_{B\in\mathcal{P}(A)} \in \bigoplus_{B\in\mathcal{P}(A)} H_B$. Then for all $\pi \in \Gamma(A)$,*

$$\sum_{B\in\mathcal{P}(A)} \phi_A F_B(\pi) = \sum_{1\le i<j\le k} \frac{F_{c(\pi_{[\![i,j]\!]})}\left(\pi_{[\![i,j]\!]}\right)}{(k-j+i)!}.$$

Proof. By definition of the embedding operator,

$$\begin{aligned}\phi_A F_B(\pi) &= \sum_{\pi'\in\Gamma(B)} F_B(\pi') \frac{\mathbb{I}\{\pi' \sqsubset \pi\}}{(k-|\pi'|+1)!} \\ &= F_B(\pi_{|B}) \frac{\mathbb{I}\{\pi_{|B} \sqsubset \pi\}}{(k-|B|+1)!}.\end{aligned}$$

Thus only the terms $\phi_A F_B(\pi)$ where B is such that $\pi_{|B}$ is a contiguous subword of π are potentially not null in the sum $\sum_{B\in\mathcal{P}(A)} \phi_A F_B(\pi)$. As the contiguous subwords of π are all of the form $\pi_{[\![i,j]\!]}$ with $1 \le i < j \le k$, this concludes the proof. ☐

3. Technical proofs of Section 4

For a random variable $\mathbf{X}$ in $\mathbb{R}^d$ with $d \ge 1$ and a sigma-algebra $\mathcal{B}$, we denote by $\mathbb{E}[\mathbf{X}|\mathcal{B}]$ the conditional expectation of $\mathbf{X}$ given $\mathcal{B}$, and define $\mathbb{V}ar[\mathbf{X}|\mathcal{B}] := \mathbb{E}[(\mathbf{X} - \mathbb{E}[\mathbf{X}])^2|\mathcal{B}]$.

Proof of Proposition 1. Since $\sum_{\sigma\in\mathfrak{S}_n} \widehat{q}_N(\sigma) = 1$, one has for any $A \in \mathcal{P}([\![n]\!])$

$$\begin{aligned}\|M_A \widehat{q}_N - P_A\|_A^2 &= \Big\| \sum_{B\in\mathcal{P}(A)} \phi_A (X_B \widehat{q}_N - X_B p) \Big\|_A^2 \\ &\le 2^{|A|} \sum_{B\in\mathcal{P}(A)} \|\phi_A(X_B \widehat{q}_N - X_B p)\|_A^2,\end{aligned}$$

using Theorem 3 and the Cauchy-Schwarz inequality. For $A, B \in \mathcal{P}([\![n]\!])$ with $A \subset B$, $F \in L(\Gamma(B))$ and $\pi \in$

$$
\begin{cases}
q_{1234} + q_{1243} + q_{1324} + q_{1342} + q_{1423} + q_{1432} + q_{2134} + q_{2143} + q_{2413} + q_{4123} + q_{4132} + q_{4213} = P_{\{1,3\}}(13) \\
q_{2314} + q_{2341} + q_{2431} + q_{3124} + q_{3142} + q_{3214} + q_{3241} + q_{3412} + q_{3421} + q_{4231} + q_{4312} + q_{4321} = P_{\{1,3\}}(31) \\
q_{1234} + q_{1243} + q_{1324} + q_{2134} + q_{2143} + q_{2314} + q_{2341} + q_{2413} + q_{2431} + q_{3124} + q_{3212} + q_{3241} = P_{\{2,4\}}(24) \\
q_{1342} + q_{1423} + q_{1423} + q_{3142} + q_{3412} + q_{3421} + q_{4123} + q_{4132} + q_{4213} + q_{4231} + q_{4312} + q_{4321} = P_{\{2,4\}}(42) \\
q_{1234} + q_{1324} + q_{1342} + q_{2134} + q_{2314} + q_{2341} + q_{3124} + q_{3142} + q_{3214} + q_{3241} + q_{3412} + q_{3421} = P_{\{3,4\}}(34) \\
q_{1243} + q_{1423} + q_{1432} + q_{2143} + q_{2413} + q_{2431} + q_{4123} + q_{4132} + q_{4213} + q_{4231} + q_{4312} + q_{4321} = P_{\{3,4\}}(43) \\
q_{1234} + q_{1243} + q_{1423} + q_{4123} = P_{\{1,2,3\}}(123) \\
q_{1324} + q_{1342} + q_{1432} + q_{4132} = P_{\{1,2,3\}}(132) \\
q_{2134} + q_{2143} + q_{2413} + q_{4213} = P_{\{1,2,3\}}(213) \\
q_{2314} + q_{2341} + q_{2431} + q_{4231} = P_{\{1,2,3\}}(231) \\
q_{3124} + q_{3142} + q_{3412} + q_{4312} = P_{\{1,2,3\}}(312) \\
q_{3214} + q_{3241} + q_{3421} + q_{4321} = P_{\{1,2,3\}}(321) \\
q_{1234} + q_{1324} + q_{1342} + q_{2134} = P_{\{1,3,4\}}(134) \\
q_{1243} + q_{1423} + q_{1432} + q_{2143} = P_{\{1,3,4\}}(143) \\
q_{2314} + q_{3124} + q_{3142} + q_{3214} = P_{\{1,3,4\}}(314) \\
q_{2341} + q_{3241} + q_{3412} + q_{3421} = P_{\{1,3,4\}}(341) \\
q_{2413} + q_{4123} + q_{4132} + q_{4213} = P_{\{1,3,4\}}(413) \\
q_{2431} + q_{4231} + q_{4312} + q_{4321} = P_{\{1,3,4\}}(431)
\end{cases}
$$

Figure 1. Linear system to find a function q on $\mathfrak{S}_4$ with the same marginals as p for $A \in \{\{1,3\},\{2,4\},\{3,4\},\{1,2,3\},\{1,3,4\}\}$

$\Gamma(A)$, one has by definition of the embedding operator, $\phi_A F(\pi) = F(\pi_{|B})\mathbb{I}\{\pi_{|B} \sqsubset \pi\}/(|A|-|B|+1)!$ and thus

$$
\begin{aligned}
\|\phi_A F\|_A^2 &= \sum_{\pi\in\Gamma(A)} \frac{F(\pi_{|B})^2}{(|A|-|B|+1)!^2}\mathbb{I}\{\pi_{|B} \sqsubset \pi\} \\
&= \sum_{\pi'\in\Gamma(B)} \frac{F(\pi')^2}{(|A|-|B|+1)!^2} \sum_{\substack{\pi\in\Gamma(A) \\ \pi_{|B}=\pi'}} \mathbb{I}\{\pi_{|B} \sqsubset \pi\}.
\end{aligned}
$$

Now, for $\pi \in \Gamma(A)$ and $\pi' \in \Gamma(B)$, $\pi' = \pi_{|B}$ and $\pi_{|B} \sqsubset \pi$ is equivalent to $\pi' \sqsubset \pi$, so $\sum_{\pi\in\Gamma(A),\, \pi_{|B}=\pi'} \mathbb{I}\{\pi_{|B} \sqsubset \pi\} = |\{\pi \in \Gamma(A) \mid \pi' \sqsubset \pi\}|$. It is easy to see that this last value is equal to the number of permutations on the set $(c(\pi)\backslash c(\pi'))\cup\{\diamond\}$ where $\diamond$ is an element that represents the block π'. It is thus equal to $(|A|-|B|+1)!$, and therefore $\|\phi_A X\|_A^2 = \|F\|_B^2/(|A|-|B|+1)!$. Injecting this result, one obtains

$$
\mathcal{E}\left(\widehat{q}_N\right) \leq \sum_{A\in\mathcal{A}} \nu(A) \sum_{B\in\mathcal{P}(A)} \frac{2^{|A|}}{(|A|-|B|+1)!} \mathbb{E}\left[\|\widehat{X}_B - X_B p\|_B^2\right].
$$

Inverting the sums concludes the proof. ☐

The proof of Proposition 2 relies on the following result.

Lemma 2. *Let $B \in \mathcal{P}(\mathcal{A})$ and $\widehat{\theta} \in \mathcal{F}(\mathcal{B}_N^{\nu}, \mathbb{R}^{2^n})$. For $\pi \in \Gamma(B)$,*

$$
\mathbb{E}\left[\widehat{X}_{B,\widehat{\theta}}(\pi)\right] = X_B p(\pi) \mathbb{E}\left[\sum_{A\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)} \widehat{\theta}(A)\right].
$$

Proof. For $\pi \in \Gamma(B)$, one has

$$
\mathbb{E}\left[\widehat{X}_{B,\widehat{\theta}}(\pi)\right] = \mathbb{E}\left[\mathbb{E}\left[\sum_{A\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)} \widehat{\theta}(A) X_B \widehat{P_A}(\pi) \,\middle|\, \mathcal{B}_N^{\nu}\right]\right].
$$

Since $\widehat{\mathcal{A}}_N$ is $\mathcal{B}_N^{\nu}$-measurable by construction and $\widehat{\theta} \in \mathcal{F}(\mathcal{B}_N^{\nu}, \mathbb{R}^{2^n})$ by hypothesis,

$$
\mathbb{E}\left[\sum_{A\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)} \widehat{\theta}(A) X_B \widehat{P_A}(\pi) \,\middle|\, \mathcal{B}_N^{\nu}\right] = \sum_{A\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)} \widehat{\theta}(A) \mathbb{E}\left[X_B \widehat{P_A}(\pi) \mid \mathcal{B}_N^{\nu}\right].
$$

Then for $A \in \widehat{\mathcal{A}}_N \cap \mathcal{Q}(B)$,

$$\begin{aligned}&\mathbb{E}\left[X_B\widehat{P_A}(\pi) \mid \mathcal{B}_N^\nu\right]\\ &= \mathbb{E}\left[\sum_{\pi'\in\Gamma(B)} \alpha_B(\pi,\pi')M_B\widehat{P_A}(\pi') \,\middle|\, \mathcal{B}_N^\nu\right]\\ &= \sum_{\pi'\in\Gamma(B)} \alpha_B(\pi,\pi') \sum_{\sigma\in\Gamma(A),\ \pi'\subset\sigma} \mathbb{E}\left[\widehat{P_A}(\sigma) \mid \mathcal{B}_N^\nu\right],\end{aligned}$$

where the $\alpha_B(\pi,\pi')$ coefficients are defined in Definition 3. Now, for $\sigma \in \Gamma(A)$, $\widehat{P_A}(\sigma) = (\sum_{i\in\widehat{I}_A} \mathbb{I}\{\pi^{(i)} = \sigma\})/|\widehat{I}_A|$ so $|\widehat{I}_A|\widehat{P_A}(\sigma) \mid \mathcal{B}_N^\nu$ is a binomial random variable of parameters $(|\widehat{I}_A|, P_A(\sigma))$, and thus $\mathbb{E}\left[\widehat{P_A}(\sigma) \mid \mathcal{B}_N^\nu\right] = P_A(\sigma)$. Therefore

$$\begin{aligned}&\mathbb{E}\left[X_B\widehat{P_A}(\pi) \mid \mathcal{B}_N^\nu\right] =\\ &\sum_{\pi'\in\Gamma(B)} \alpha_B(\pi,\pi') \sum_{\sigma\in\Gamma(A),\ \pi'\subset\sigma} P_A(\sigma) = X_BP_A(\pi),\end{aligned}$$

so that

$$\mathbb{E}\left[\widehat{X}_{B,\widehat{\theta}}(\pi)\right] = \mathbb{E}\left[\sum_{A\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)} \widehat{\theta}(A)\right] X_BP_A(\pi),$$

which is the desired result. □

Proof of Proposition 2. Using Lemma 2, one has for $A \in \mathcal{P}(\mathcal{A})$ and $\pi \in \Gamma(A)$

$$\begin{aligned}\mathbb{E}\left[M_A\widehat{q}_N(\pi)\right] &= \mathbb{E}\left[\sum_{B\in\mathcal{P}(A)\cup\{\emptyset\}} \phi_A\widehat{X}_{B,\widehat{\theta}}(\pi)\right]\\ &= \frac{1}{|A|!} + \sum_{B\in\mathcal{P}(A)} X_Bp(\pi)\mathbb{E}\left[\sum_{A'\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)} \widehat{\theta}(A')\right].\end{aligned}$$

Thus if $\lim_{n\to\infty} \mathbb{E}\left[\sum_{A\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)} \widehat{\theta}(A)\right] = 1$, then

$$M_A\widehat{q}_N(\pi) \;=\; \frac{1}{|A|!} + \sum_{B\in\mathcal{P}(A)} X_Bp(\pi) \;=\; M_Ap(\pi).$$

□

The proof of Theorem 4 relies on the two following lemmas

Lemma 3. *Let $B \in \mathcal{P}(\mathcal{A})$ and $\widehat{\theta} \in \mathcal{F}(\mathcal{B}_N^\nu, \mathbb{R}^{2^n})$. For $\pi \in \Gamma(B)$,*

$$\begin{aligned}\mathbb{V}ar\left[\widehat{X}_{B,\widehat{\theta}}(\pi)\right] &= \left(X_B^2p(\pi) - X_Bp(\pi)^2\right)\\ &\quad\times \mathbb{E}\left[\sum_{A\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)} \frac{\widehat{\theta}(A)^2}{|\widehat{I}_A|}\right],\end{aligned}$$

where X_B^2 is the operator on $L(\Gamma(B))$ defined by $X_B^2f(\pi) := \sum_{\pi'\in\Gamma(B)} \alpha_B^2(\pi,\pi')f(\pi')$ for $f \in L(\Gamma(B))$.

Proof. For $\pi \in \Gamma(B)$, one has

$$\begin{aligned}&\mathbb{V}ar\left[\widehat{X}_{B,\widehat{\theta}}(\pi)\right] =\\ &\qquad\mathbb{E}\left[\mathbb{V}ar\left[\sum_{A\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)} \widehat{\theta}(A)X_B\widehat{P_A}(\pi) \,\middle|\, \mathcal{B}_N^\nu\right]\right].\end{aligned}$$

Since $\widehat{\theta} \in \mathcal{F}(\mathcal{B}_N^\nu, \mathbb{R}^{2^n})$ by hypothesis and the $\widehat{P_A}$'s are independent conditionally to $\widehat{\mathcal{A}}_N$,

$$\begin{aligned}&\mathbb{V}ar\left[\sum_{A\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)} \widehat{\theta}(A)X_B\widehat{P_A}(\pi) \,\middle|\, \mathcal{B}_N^\nu\right] =\\ &\qquad\sum_{A\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)} \widehat{\theta}(A)^2\mathbb{V}ar\left[X_B\widehat{P_A}(\pi) \,\middle|\, \mathcal{B}_N^\nu\right].\end{aligned}$$

Now, for $A \in \mathcal{P}([\![n]\!])$,

$$\begin{aligned}\mathbb{V}ar\left[X_B\widehat{P_A}(\pi) \,\middle|\, \mathcal{B}_N^\nu\right] &= \mathbb{E}\left[\left(X_B\widehat{P_A}(\pi)\right)^2 \,\middle|\, \mathcal{B}_N^\nu\right]\\ &\quad - \mathbb{E}\left[X_B\widehat{P_A}(\pi) \,\middle|\, \mathcal{B}_N^\nu\right]^2, \quad (1)\end{aligned}$$

with

$$\begin{aligned}&X_B\widehat{P_A}(\pi)^2 =\\ &\sum_{\pi',\pi''\in\Gamma(B)} \alpha_B(\pi,\pi')\alpha_B(\pi,\pi'') \sum_{\substack{\sigma',\sigma''\in\Gamma(A)\\ \pi'\subset\sigma'\\ \pi''\subset\sigma''}} \widehat{P_A}(\sigma')\widehat{P_A}(\sigma'').\end{aligned}$$

Now, for $\sigma, \sigma' \in \Gamma(A)$,

$$\begin{aligned}&\mathbb{E}\left[\widehat{P_A}(\sigma')\widehat{P_A}(\sigma'') \,\middle|\, \mathcal{B}_N^\nu\right]\\ &= \mathbb{E}\left[\frac{1}{|\widehat{I}_A|^2} \sum_{i\in\widehat{I}_A} \mathbb{I}\{\pi^{(i)} = \sigma'\} \sum_{i\in I_A} \mathbb{I}\{\pi^{(i)} = \sigma''\} \,\middle|\, \mathcal{B}_N^\nu\right]\\ &= \frac{1}{|I_A|^2} \sum_{i,j\in I_A} \mathbb{P}\left[\pi^{(i)} = \sigma', \pi^{(j)} = \sigma''\right].\end{aligned}$$

If $i = j$,

$$\mathbb{P}\left[\pi^{(i)} = \sigma', \pi^{(j)} = \sigma''\right] = P_A(\sigma')\mathbb{I}\{\sigma' = \sigma''\}$$

and if $i \neq j$,

$$\mathbb{P}\left[\pi^{(i)} = \sigma', \pi^{(j)} = \sigma''\right] = P_A(\sigma')P_A(\sigma''),$$

because the $\pi^{(i)}$'s are independent. Thus

$$\mathbb{E}\left[\widehat{P_A}(\sigma')\widehat{P_A}(\sigma'') \,\Big|\, \mathcal{B}_N^\nu\right] = \frac{|\widehat{I}_A|-1}{|\widehat{I}_A|}P_A(\sigma')P_A(\sigma'') + \frac{\mathbb{I}\{\sigma'=\sigma''\}}{|\widehat{I}_A|}P_A(\sigma')$$

and so

$$\sum_{\substack{\sigma',\sigma''\in\Gamma(A)\\ \pi'\subset\sigma'\\ \pi''\subset\sigma''}} \mathbb{E}\left[\widehat{P_A}(\sigma')\widehat{P_A}(\sigma'') \,\Big|\, \mathcal{B}_N^\nu\right] = \frac{|\widehat{I}_A|-1}{|\widehat{I}_A|}M_BP_A(\pi')M_BP_A(\pi'') + \frac{\mathbb{I}\{\pi'=\pi''\}}{|\widehat{I}_A|}M_BP_A(\pi'),$$

because for $\pi',\pi''\in\Gamma(B)$, $\pi'\subset\sigma'$ and $\pi''\subset\sigma'$ implies $\pi'=\pi''$ (in other words, σ' has a unique subword of content B). Therefore

$$\mathbb{E}\left[\left(X_B\widehat{P_A}(\pi)\right)^2 \,\Big|\, \mathcal{B}_N^\nu\right] = \frac{|\widehat{I}_A|-1}{|\widehat{I}_A|}X_BP_A(\pi)^2 + \frac{1}{|\widehat{I}_A|}\sum_{\pi'\in\Gamma(B)}\alpha_B(\pi,\pi')^2M_BP_A(\pi')$$

and injecting this result into (1) gives

$$\mathbb{V}ar\left[X_B\widehat{P_A}(\pi) \,\Big|\, \mathcal{B}_N^\nu\right] = \frac{1}{|\widehat{I}_A|}\left(X_B^2P_A(\pi) - X_BP_A(\pi)^2\right),$$

where $X_B^2 f := \sum_{\pi'\in\Gamma(B)}\alpha_B(\pi,\pi')^2M_Bf(\pi')$ for $f\in\bigsqcup_{A\in\mathcal{P}([\![n]\!])}L(\Gamma(A))$. Gathering all the calculations, one obtains

$$\mathbb{V}ar\left[\widehat{X}_{B,\widehat{\theta}}(\pi)\right] = \mathbb{E}\left[\sum_{A\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)}\frac{\widehat{\theta}(A)^2}{|\widehat{I}_A|}\right] \times \left(X_B^2P_A(\pi) - X_BP_A(\pi)^2\right).$$

□

Lemma 4. *For all $B\in\mathcal{P}(\mathcal{A})$,*

$$\mathbb{E}\left[\sum_{A\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)}\widehat{\theta}^{WLS}(A)\right] = 1-(1-\nu[\mathcal{Q}(B)])^N,$$

$$\mathbb{E}\left[\sum_{A\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)}\frac{\widehat{\theta}^{WLS}(A)^2}{|\widehat{I}_A|}\right] = \mathbb{E}\left[\frac{\mathbb{I}\{Z_N^B\geq 1\}}{Z_N^B}\right],$$

where for any collection $\mathcal{S}\subset\mathcal{P}([\![n]\!])$, $\nu[\mathcal{S}] := \sum_{A\in\mathcal{S}}\nu(A)$, and Z_N^B is a binomial random variable of parameters N and $\nu[\mathcal{Q}(B)]$.

Proof. By definition, the coefficients of the WLS estimator are given for all $A\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)$ by

$$\widehat{\theta}^{WLS}(A) := \frac{\widehat{\nu}_N(A)}{\sum_{A'\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)}\widehat{\nu}_N(A')}.$$

For $B\in\mathcal{P}(A)$,

$$\begin{aligned}\mathbb{E}&\left[\sum_{A\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)}\widehat{\theta}^{WLS}(A)\right]\\ &= \mathbb{E}\left[\sum_{A\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)}\widehat{\theta}^{WLS}(A)\middle| B\in\mathcal{P}(\widehat{\mathcal{A}}_N)\right] \times \mathbb{P}\left[B\in\mathcal{P}(\widehat{\mathcal{A}}_N)\right],\end{aligned}$$

because $\sum_{A\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)}\widehat{\theta}^{WLS}(A) = 0$ when $\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)=\emptyset$. On the one hand,

$$\begin{aligned}\mathbb{E}&\left[\sum_{A\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)}\widehat{\theta}^{WLS}(A)\middle| B\in\mathcal{P}(\widehat{\mathcal{A}}_N)\right]\\ &= \mathbb{E}\left[\sum_{A\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)}\frac{\widehat{\nu}_N(A)}{\sum_{A'\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)}\widehat{\nu}_N(A')}\right] = 1,\end{aligned}$$

and on the other hand,

$$\begin{aligned}\mathbb{P}\left[B\in\mathcal{P}(\widehat{\mathcal{A}}_N)\right] &= 1-\mathbb{P}\left[B\notin\mathcal{P}(\widehat{\mathcal{A}}_N)\right]\\ &= 1-\mathbb{P}\left[\bigcap_{i=1}^N\{B\not\subset A_i\}\right]\\ &= 1-\mathbb{P}\left[B\not\subset\mathbf{A}\right]^N,\end{aligned}$$

where $\mathbf{A}$ is a random variable on $\mathcal{P}([\![n]\!])$ of law ν. Then $\mathbb{P}[B\not\subset\mathbf{A}] = 1-\mathbb{P}[B\subset\mathbf{A}] = 1-\sum_{A\in\mathcal{Q}(B)}\nu(A)$.

Similarly,

$$\begin{aligned}\mathbb{E}&\left[\sum_{A\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)}\frac{\widehat{\theta}^{WLS}(A)^2}{|\widehat{I}_A|}\right]\\ &= \mathbb{E}\left[\sum_{A\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)}\frac{\widehat{\theta}^{WLS}(A)^2}{|\widehat{I}_A|}\mathbb{I}\{B\in\mathcal{P}(\widehat{\mathcal{A}}_N)\}\right]\\ &= \mathbb{E}\left[\sum_{A\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)}\frac{\widehat{\nu}_N(A)^2\mathbb{I}\{B\in\mathcal{P}(\widehat{\mathcal{A}}_N)\}}{|\widehat{I}_A|(\sum_{A'\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)}\widehat{\nu}_N(A'))^2}\right]\\ &= \mathbb{E}\left[\frac{\sum_{A\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)}|\widehat{I}_A|}{(\sum_{A'\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)}|\widehat{I}_{A'}|)^2}\mathbb{I}\{B\in\mathcal{P}(\widehat{\mathcal{A}}_N)\}\right]\\ &= \mathbb{E}\left[\frac{1}{\sum_{A\in\widehat{\mathcal{A}}_N\cap\mathcal{Q}(B)}|\widehat{I}_A|}\mathbb{I}\{B\in\mathcal{P}(\widehat{\mathcal{A}}_N)\}\right]\end{aligned}$$

(we recall that $\widehat{\nu}_N(A) = |\widehat{I}_A|/N$). Now, by definition

$$\sum_{A \in \widehat{\mathcal{A}}_N \cap \mathcal{Q}(B)} |\widehat{I}_A| = \sum_{i=1}^{N} \mathbb{I}\{A_i \supset B\}.$$

We denote by Z_B this random variable. The A_i's being IID, Z_N^B is a binomial random variable of parameters N and $\mathbb{P}[A_i \supset B] = \nu[\mathcal{Q}(B)]$. Furthermore, the event $\{B \in \mathcal{P}(\widehat{\mathcal{A}}_N)\}$ is equal to $\{Z_N^B \geq 1\}$. Thus in conclusion,

$$\mathbb{E}\left[\sum_{A \in \widehat{\mathcal{A}}_N \cap \mathcal{Q}(B)} \frac{\widehat{\theta}^{WLS}(A)^2}{|\widehat{I}_A|}\right] = \mathbb{E}\left[\frac{\mathbb{I}\{Z_N^B \geq 1\}}{Z_N^B}\right].$$

□

Proof of Theorem 4. For $B \in \mathcal{P}([\![n]\!])$ and $\widehat{X}_B \in \mathcal{F}(\mathcal{B}_N, L(\Gamma(B)))$, one has the usual bias-variance decomposition

$$\mathbb{E}\left[\|\widehat{X}_B - X_B p\|_B^2\right] = \sum_{\pi \in \Gamma(B)} \left(\mathbb{E}\left[\widehat{X}_B(\pi)\right] - X_B p(\pi)\right)^2 + \sum_{\pi \in \Gamma(B)} \mathbb{V}ar\left[\widehat{X}_B(\pi)\right].$$

Therefore, combining Proposition 1, and Lemmas 3 and 4, one obtains

$$\begin{aligned}
\mathcal{E}\left(\widehat{q}_N^{WLS}\right) &\leq \sum_{B \in \mathcal{P}(\mathcal{A})} \nu_\phi(B) \mathbb{E}\left[\|\widehat{X}_B^{WLS} - X_B p\|_B^2\right] \\
&\leq \sum_{B \in \mathcal{P}(\mathcal{A})} \nu_\phi(B) \left(\sum_{\pi \in \Gamma(B)} \left(\mathbb{E}\left[\widehat{X}_B^{WLS}(\pi)\right] - X_B p(\pi)\right)^2 + \sum_{\pi \in \Gamma(B)} \mathbb{V}ar\left[\widehat{X}_B^{WLS}(\pi)\right]\right) \\
&\leq \sum_{B \in \mathcal{P}(\mathcal{A})} \nu_\phi(B) \left(\|X_B p\|_B^2 (1 - \nu[\mathcal{Q}(B)])^{2N} + \sum_{\pi \in \Gamma(B)} \left(X_B^2 p(\pi) - X_B p(\pi)^2\right) \mathbb{E}\left[\frac{\mathbb{I}\{Z_N^B \geq 1\}}{Z_N^B}\right]\right).
\end{aligned}$$

Notice that for $z \geq 1$, $z + 1 \leq 2z$, so that

$$\frac{\mathbb{I}\{Z_N^B \geq 1\}}{Z_N^B} \leq \frac{2}{Z_N^B + 1}.$$

Now, Chao & Strawderman (1972) provides the following closed-form expression, for a binomial random variable Z of parameters (n, p),

$$\mathbb{E}\left[\frac{1}{Z+1}\right] = \frac{1 - (1-p)^{n+1}}{p(n+1)}.$$

Therefore,

$$\mathbb{E}\left[\frac{\mathbb{I}\{Z_N^B \geq 1\}}{Z_N^B}\right] \leq \frac{2}{\nu[\mathcal{Q}(B)](N+1)}.$$

Defining the constants

$$\begin{aligned}
C_1 &= 2 \sum_{B \in \mathcal{P}(\mathcal{A})} \frac{\nu_\phi(B) \sum_{\pi \in \Gamma(B)} \left(X_B^2 p(\pi) - X_B p(\pi)^2\right)}{\nu[\mathcal{Q}(B)]} \\
C_2 &= \sum_{B \in \mathcal{P}(\mathcal{A})} \nu_\phi(B) \|X_B p\|_B^2 \\
\rho &= 1 - \min_{B \in \mathcal{P}(\mathcal{A})} \nu[\mathcal{Q}(B)]
\end{aligned}$$

gives the desired formula. Since $\mathcal{Q}(B) \cap \mathcal{A} \neq \emptyset$ for $B \in \mathcal{P}(\mathcal{A})$, one has $\nu[\mathcal{Q}(B)] > 0$ for all $B \in \mathcal{P}(\mathcal{A})$ and therefore $\rho < 1$. This concludes the proof. □

4. Computation of wavelet projections

The computation of wavelet projections only involves the parameters $\alpha_B(\pi, \pi')$ for $B \in \mathcal{P}([\![n]\!])$ and $\pi, \pi' \in \Gamma(B)$. Their computation can be made once and for all applications. Here we show how to perform it efficiently. The first simplification comes from the following lemma, established in Clémençon et al. (2014). For $\pi = \pi_A \ldots \pi_k \in \Gamma_n$ and $\sigma \in \mathfrak{S}_n$, we denote by $\sigma(\pi)$ the word $\sigma(\pi_1) \ldots \sigma(\pi_k) \in \Gamma(\sigma(c(\pi)))$.

Lemma 5. *Let $B \in \mathcal{P}([\![n]\!])$ and $\sigma \in \mathfrak{S}_n$ a permutation that keeps the order of the items in B, i.e. such that for all $b, b' \in B$, $b < b' \Rightarrow \sigma(b) < \sigma(b')$. Then for all $\pi, \pi' \in \Gamma(B)$,*

$$\alpha_B(\pi, \pi') = \alpha_{\sigma(B)}(\sigma(\pi), \sigma(\pi')).$$

Lemma 5 implies two simplifications:

- First, for $k \in \{2, \ldots, n\}$, the coefficients $(\alpha_B(\pi, \pi'))_{\pi, \pi' \in \Gamma(B)}$ are obtained directly from the $(\alpha_{\{1,\ldots,k\}}(\pi, \pi'))_{\pi, \pi' \in \Gamma(\{1,\ldots,k\})}$ for all $B \subset [\![n]\!]$ with $|B| = k$.
- Second, for $B = \{b_1, \ldots, b_k\} \in \mathcal{P}([\![n]\!])$ with $b_1 < \cdots < b_k$, the coefficients $(\alpha_B(\pi, \pi'))_{\pi' \in \Gamma(B)}$ are obtained directly from the $(\alpha_B(b_1 \ldots b_k, \pi'))_{\pi' \in \Gamma(B)}$ for any $\pi \in \Gamma(B)$.

Example 1. Let $B = \{2, 4, 5\}$ and $\sigma \in \mathfrak{S}_n$ such that $\sigma(2) = 1, \sigma(4) = 2$ and $\sigma(5) = 3$. Then for $\pi, \pi' \in \Gamma(\{2, 4, 5\})$, $\alpha_{\{2,4,5\}}(\pi, \pi') = \alpha_{\{1,2,3\}}(\sigma(\pi), \sigma(\pi'))$.

With the precedent simplifications, one only needs to compute and store the $k!$ coefficients $(\alpha_{\{1,\ldots,k\}}(12 \ldots k, \pi))_{\pi \in \Gamma(\{1,\ldots,k\})}$ for each $k \in \{2, \ldots, K\}$. We now further describe how the computation of each $\alpha_{\{1,\ldots,k\}}(12 \ldots k, \pi)$ can be made efficiently. By construction

$$X_A F = F - \phi_A X_\emptyset F - \sum_{B \in \mathcal{P}(A) \setminus \{A\}} \phi_A X_B F$$

for any $A \in \mathcal{P}([\![n]\!])$ and $F \in L(\Gamma(A))$. Applying Lemma 1 gives the following recursive formula for the $\alpha_B(\pi, \pi')$'s.

Lemma 6. *Let $B \in \mathcal{P}([\![n]\!])$ and $k = |B|$. Then for all $\pi, \pi' \in \Gamma(B)$,*

$$\alpha_B(\pi, \pi') = \mathbb{I}\{\pi = \pi'\} - \frac{1}{k!} - \sum_{\substack{1 \le i < j \le k \\ j - i < k-1}} \frac{1}{(k-j+i)!} \alpha_{c(\pi_{[\![i,j]\!]})} \left(\pi_{[\![i,j]\!]}, \pi'_{|c(\pi_{[\![i,j]\!]})} \right).$$

Using lemma 6, it is easy to see that the computation of all the $\alpha_{\{1,\dots,k\}}(12\dots k, \pi)$ for $\pi \in \Gamma(\{1,\dots,k\})$ and $k \in \{2,\dots,K\}$ can be implemented with complexity bounded by k^2. Combined with all the precedent simplifications, this shows the following result.

Lemma 7. *For $K \in \{2,\dots,n\}$, the computation of all coefficients $\alpha_B(\pi, \pi')$ for $\pi, \pi' \in \Gamma(B)$ and $B \in \mathcal{P}([\![n]\!])$ with $|B| \le K$ has complexity bounded by $K^2 K!$.*

Example 2. The following tables give the values of the coefficients $(\alpha_{\{1,\dots,k\}}(12\dots k, \pi))_{\pi \in \Gamma(\{1,\dots,k\})}$ for $k = 2$:

π	$\alpha_{\{1,2\}}(12, \pi)$
12	1/2
21	−1/2

and $k = 3$:

π	$\alpha_{\{1,2,3\}}(123, \pi)$
123	1/3
132	−1/6
213	−1/6
231	−1/6
312	−1/6
321	1/3

5. Technical proofs of Section 5

Proof of Proposition 3. By construction, any MRA-based linear ranking model can be stored directly as the collection of estimators $(\widehat{X}_B)_{B \in \mathcal{P}(\widehat{\mathcal{A}}_N)}$ and not as a function on $\mathfrak{S}_n$. Denoting by $\widehat{\mathcal{N}}$ the total number of parameters to be stored, one then has

$$\widehat{\mathcal{N}} \le \sum_{B \in \mathcal{P}(\widehat{\mathcal{A}}_N)} |B|! \le K!\, |\mathcal{P}(\widehat{\mathcal{A}}_N)| \le K! \sum_{A \in \widehat{\mathcal{A}}_N} 2^{|A|} \le K!\, 2^K |\widehat{\mathcal{A}}_N|,$$

which gives the result because $|\widehat{\mathcal{A}}_N| \le \min(N, |\mathcal{A}|)$. Now for any $A \in \mathcal{P}([\![n]\!])$, the marginal on A of $\widehat{q}_N$ is given by $M_A \widehat{q}_N = \sum_{B \in \mathcal{P}(A) \cup \{\emptyset\}} \phi_A \widehat{X}_B$, where we set by convention $\widehat{X}_B = 0$ for $B \in \mathcal{P}([\![n]\!]) \setminus \mathcal{P}(\mathcal{A})$. Applying Lemma 1, one then has for any $\pi \in \Gamma(A)$ with $k = |A|$,

$$M_A \widehat{q}_N(\pi) = \frac{1}{k!} + \sum_{1 \le i < j \le k} \frac{1}{(k-j+i)!} \widehat{X}_{c(\pi_{[\![i,j]\!]})}(\pi_{[\![i,j]\!]}).$$

The computation of $M_A \widehat{q}_N(\pi)$ thus requires at most $k(k-1)/2$ operations. □

Proof of Proposition 4. Using the formula of Definition 4,

$$\begin{aligned}
\widehat{X}_B^{WLS}(\pi) &= \sum_{A \in \widehat{\mathcal{A}}_N \cap \mathcal{Q}(B)} \widehat{\theta}^{WLS}(A) X_B \widehat{P_A}(\pi) \\
&= \sum_{A \in \widehat{\mathcal{A}}_N \cap \mathcal{Q}(B)} \frac{\widehat{\nu}_N(A)}{\sum_{A' \in \widehat{\mathcal{A}}_N \cap \mathcal{Q}(B)} \widehat{\nu}_N(A')} \sum_{\pi' \in \Gamma(B)} \alpha_B(\pi, \pi') M_B \widehat{P_A}(\pi').
\end{aligned}$$

Now, for $A \in \mathcal{P}([\![n]\!])$ and $\pi' \in \Gamma(B)$,

$$\begin{aligned}
M_B \widehat{P_A}(\pi') &= \sum_{\pi \in \Gamma(A),\ \pi' \subset \pi} \widehat{P_A}(\pi) \\
&= \sum_{\pi \in \Gamma(A),\ \pi' \subset \pi} \frac{1}{|\widehat{I}_A|} \sum_{i \in \widehat{I}_A} \mathbb{I}\{\pi^{(i)} = \pi\} \\
&= \frac{1}{|\widehat{I}_A|} |\{i \in \widehat{I}_A \mid \pi' \subset \pi\}|.
\end{aligned}$$

Thus, recalling that $\widehat{\nu}_N(A) = |\widehat{I}_A|/N$ for $A \in \mathcal{P}([\![n]\!])$,

$$\widehat{X}_B^{WLS}(\pi) = \frac{1}{\sum_{A' \in \widehat{\mathcal{A}}_N \cap \mathcal{Q}(B)} |\widehat{I}_{A'}|} \times \sum_{A \in \widehat{\mathcal{A}}_N \cap \mathcal{Q}(B)} |\{i \in \widehat{I}_A \mid \pi' \subset \pi\}|,$$

which concludes the proof. □

Proof of Proposition 5. The explicit formula given by Proposition 4 for the WLS estimators $\widehat{X}_B^{WLS}$ shows that their computation can be decomposed in two steps:

- Compute all the $|\widehat{I}_A|$ for $A \in \widehat{\mathcal{A}}_N$ and all the $|\{1 \le i \le N \mid \pi' \subset \pi^{(i)}\}|$ for $\pi \in \Gamma_n$ such that $c(\pi) \in \mathcal{P}(\widehat{\mathcal{A}}_N)$
- Compute all the $\widehat{X}_B^{WLS}(\pi)$ for all $B \in \mathcal{P}(\widehat{\mathcal{A}}_N)$ and $\pi \in \Gamma(B)$ using the quantities computed in the first step and the pre-computed coefficients $\alpha_B(\pi, \pi')$ for $\pi, \pi' \in \Gamma(B)$ and $B \in \mathcal{P}(\widehat{\mathcal{A}}_N)$.

The first step is a simple counting of occurrence numbers and it can be performed in one loop over the dataset $\mathcal{D}_N$ with complexity bounded by $N2^K$. The second step requires for each couple $(B, \pi) \in \mathcal{P}(\widehat{\mathcal{A}}_N) \times \Gamma(B)$ at most $\sum_{A \in \widehat{\mathcal{A}}_N \cap \mathcal{Q}(B)} (|\widehat{I}_A| + 1)$ operations. Indeed, the number of rankings $\pi' \in \Gamma(B)$ for which $|\{1 \leq i \leq N \mid \pi' \subset \pi^{(i)}\}| \neq 0$ is bounded by $\sum_{A \in \widehat{\mathcal{A}}_N \cap \mathcal{Q}(B)} |\widehat{I}_A|$. The global complexity of the second step is therefore bounded by

$$\begin{aligned}
\sum_{B \in \mathcal{P}(\widehat{\mathcal{A}}_N)} |B|! \sum_{A \in \widehat{\mathcal{A}}_N \cap \mathcal{Q}(B)} (|\widehat{I}_A| + 1) \\
&\leq K! \sum_{A \in \widehat{\mathcal{A}}_N} \sum_{B \in \mathcal{P}(A)} (|\widehat{I}_A| + 1) \\
&\leq K! 2^K \sum_{A \in \widehat{\mathcal{A}}_N} (|\widehat{I}_A| + 1) \\
&\leq K! 2^K (N + |\mathcal{A}|),
\end{aligned}$$

because $\sum_{A \in \widehat{\mathcal{A}}_N} |\widehat{I}_A| = N$ by definition and $|\widehat{\mathcal{A}}_N| \leq |\mathcal{A}|$. The global complexity of the two steps is then bounded by $2^K (K! + 1)(N + |\mathcal{A}|)$. $\square$

References

Chao, M.T. and Strawderman, W.E. Negative moments of positive random variables. *Journal of the American Statistical Society*, 67:429–431, 1972.

Clémençon, Stéphan, Jakubowicz, Jérémie, and Sibony, Eric. Multiresolution analysis of incomplete rankings. *ArXiv e-prints*, march 2014.