

Maximum de la distance de transfert à une partition donnée

Maximum of the transfer distance to a given partition

Irène Charon, Lucile Denoeud et Olivier Hudry



Édition électronique

URL : <http://journals.openedition.org/msh/7403>

DOI : 10.4000/msh.7403

ISSN : 1950-6821

Éditeur

Centre d'analyse et de mathématique sociales de l'EHESS

Édition imprimée

Date de publication : 1 septembre 2007

Pagination : 45-83

ISSN : 0987-6936

Référence électronique

Irène Charon, Lucile Denoeud et Olivier Hudry, « Maximum de la distance de transfert à une partition donnée », *Mathématiques et sciences humaines* [En ligne], 179 | Automne 2007, mis en ligne le 21 décembre 2007, consulté le 23 juillet 2020. URL : <http://journals.openedition.org/msh/7403> ; DOI : <https://doi.org/10.4000/msh.7403>

MAXIMUM DE LA DISTANCE DE TRANSFERT À UNE PARTITION DONNÉE

Irène CHARON¹, Lucile DENCEUD^{1,2}, Olivier HUDRY^{1,2}

RÉSUMÉ – *Dans cet article, on étudie une distance définie sur l'ensemble des partitions d'un ensemble fini. Étant données deux partitions \mathcal{P} et \mathcal{Q} , cette distance, appelée distance de transfert, correspond au nombre minimum de transferts d'un élément d'une classe dans une autre nécessaires pour transformer \mathcal{P} en \mathcal{Q} , ou réciproquement \mathcal{Q} en \mathcal{P} . On détermine le maximum de la distance de transfert entre une partition donnée et une partition quelconque, indépendamment du nombre de classes, puis le maximum de cette distance entre une partition donnée et une partition dont le nombre de classes est majoré.*

MOTS CLÉS – Distance, Partitions, Transferts

SUMMARY – *Maximum of the transfer distance to a given partition*
In this paper, we study a distance defined over the partitions of a finite set. Given two partitions \mathcal{P} and \mathcal{Q} , this distance, called the transfer distance, is defined as the minimum number of transfers of an element from one class to another, required to transform \mathcal{P} into \mathcal{Q} or equivalently \mathcal{Q} in \mathcal{P} . We give some formulae for the maximum distance value between a given partition and any partition, then the maximum distance value between a given partition and a partition with an upper-bounded number of classes.

KEYWORDS – Distance, Partitions, Transfers

1. INTRODUCTION

Le partitionnement de grands ensembles est un sujet d'étude dans plusieurs domaines d'une grande importance tels que :

- les réseaux sociaux [Seidman, 1983], dans lesquels on cherche à découvrir des communautés, que ce soit à travers la relation de citation entre auteurs de publications [Batagelj, Mrvar, Zaversnik, 1999 ; Newman, 2001], ou la relation de liens entre pages Web [Moody, 2001],
- la pagination des circuits électroniques, pour leur conception et leur représentation [Alpert, Kang, 1995 ; Fraysseix (de), Kuntz, 1992],

¹École Nationale Supérieure des Télécommunications (ENST), 46 rue Barrault 75634 Paris cedex 13, charon@infres.enst.fr

²Centre de Recherche de Mathématiques Statistiques et Économie (CERMSEM), Université Paris 1 - Panthéon-Sorbonne, 106-112 bd de l'Hôpital 75647 Paris cedex 13, denoed@infres.enst.fr, hudry@infres.enst.fr

- la biologie moléculaire, en particulier pour la classification des protéines, que ce soit à partir des données d'expression [Getz, Levine, Domany, 2000] ou des graphes d'interactions directes entre protéines [Guénoche, 2004 ; Matsuda, Ishihara, Hashimoto, 1999 ; van Dongen, 2000].

À chaque fois, les ensembles que l'on cherche à partitionner sont de grande taille, plusieurs milliers à plusieurs dizaines de milliers d'éléments. Dans chacun de ces domaines, les chercheurs conçoivent et développent de nouvelles méthodes de partitionnement. Se pose alors le problème de valider ces méthodes ou de les comparer les unes aux autres.

Un protocole simple, pour évaluer une méthode de partitionnement, consiste à mesurer son aptitude à retrouver des partitions initialement mises dans les données. On dispose d'un ensemble X , avec une partition naturelle \mathcal{P} . En appliquant une quelconque méthode de partitionnement, on obtient une partition \mathcal{Q} que l'on veut comparer à \mathcal{P} et qui n'a pas nécessairement le même nombre de classes. On peut aussi vouloir comparer directement deux partitions \mathcal{Q} et \mathcal{Q}' obtenues par deux méthodes différentes à partir des mêmes données.

Le problème n'est pas nouveau [Arabie, Hubert, 1996 ; Hubert, Arabie, 1985 ; Lerman, 1988 ; Rand, 1971], et l'article de Day [1981] établit une liste de dix distances entre partitions, qu'il compare du point de vue de la complexité de leur évaluation et d'une relation d'ordre partiel (deux métriques sont comparables si les valeurs de l'une sont systématiquement inférieures ou égales aux valeurs de l'autre). Plusieurs de ces distances sont définies comme le plus petit nombre de transformations sur les classes pour passer de \mathcal{P} à \mathcal{Q} et sont baptisées *Minimum Length Sequence metrics* par Day ; ou encore *distances d'édition*, les transformations invoquées étant des opérations d'édition sur les partitions (ajout et retrait d'un élément, fusion et fission de classes). Nous étudions dans cet article celle qui correspond à la seule opération de transfert d'un élément d'une classe dans une autre, éventuellement vide, et qu'on appellera *distance de transfert* (notée par la suite θ).

La distance de transfert fut tout d'abord introduite par Régnier en 1965 [Régnier, 1983]. En 1981, Day [1981] montre que le calcul de cette distance revient à la résolution d'un problème de couplage maximum de poids minimum dans un graphe biparti. On peut trouver une application de cette distance dans [Guénoche, 2005], où on l'utilise pour comparer différents algorithmes de partitionnement de graphe. Elle est également étudiée dans [Dencéud, Guénoche, 2006] et sert à comparer des indices usuels de distance entre partitions.

Cette distance présente cependant le défaut de ne pas être normalisée. Cet article propose des valeurs extrémales de la distance de transfert sur différents ensembles de partitions, ce qui permet de normaliser celle-là sur ceux-ci. Il est trivial de trouver le maximum θ_{\max} de la distance de transfert entre deux partitions quelconques sur un ensemble X ; il correspond au diamètre de l'ensemble des partitions sur X : $\theta_{\max} = |X| - 1$ (c'est la distance entre la partition à une classe et la partition à $|X|$ classes).

Si on veut comparer des partitions dont certaines caractéristiques sont fixées, on peut alors vouloir évaluer le maximum de la distance de transfert sur ce nouvel

ensemble de partitions afin d'obtenir une normalisation plus adaptée. Dans [Charon, Dencœud, Guénoche, Hudry, 2006], nous avons déterminé la distance maximum entre deux partitions dont les nombres de classes sont donnés ou majorés. Nous poursuivons cette étude dans cet article en établissant le maximum de la distance de transfert à une partition \mathcal{P} donnée et le maximum de la distance de transfert entre une partition \mathcal{P} donnée et une partition dont le nombre de classes est donné ou majoré.

L'exemple suivant illustre une telle situation. Une chaîne de télévision veut, à l'approche d'une élection, organiser des émissions politiques où figureraient des représentants des différents partis en compétition. Le nombre d'émissions, q , peut être fixé, majoré ou inconnu. La chaîne décide d'inviter n représentants de partis politiques ; soit X l'ensemble des représentants ; le regroupement des n représentants selon leur appartenance politique définit une partition \mathcal{P} de X . Pour des raisons de diversité, chaque représentant ne participe qu'à une seule émission ; en revanche, on ne s'impose pas que tous les partis soient représentés à chaque fois. L'ensemble des émissions constitue donc une partition \mathcal{Q} de X , le nombre de classes de \mathcal{Q} étant le nombre q d'émissions à programmer. Une façon de constituer les q émissions tout en cherchant à varier leur composition consiste à répartir les n représentants en une partition \mathcal{Q} maximisant la distance de transfert $\theta(\mathcal{P}, \mathcal{Q})$ par rapport à \mathcal{P} . On retrouve alors les problèmes étudiés ici, avec les variantes liées au fait que q est fixé, majoré ou totalement libre.

De façon plus générale, on retrouve ce problème dès que l'on cherche à répartir des éléments formant des classes homogènes de manière à créer d'autres classes, au contraire hétérogènes. C'est le cas par exemple pour la constitution de jurys, de commissions d'experts, d'équipes de travail, etc., à compétences multiples et complémentaires : X est l'ensemble des n personnes concernées ; le regroupement de ces personnes selon leur domaine de compétence forme une partition \mathcal{P} de X ; la multiplicité de compétence des jurys, commissions, équipes... est obtenue en mélangeant les personnes selon une partition \mathcal{Q} maximisant la distance de transfert à \mathcal{P} .

Une dernière illustration se rencontre par exemple dans le cadre d'un week-end d'intégration concernant de nouveaux élèves d'un établissement scolaire. L'objectif est ici d'organiser des activités (sportives, ludiques ou plus sérieuses) en groupes de façon à favoriser le brassage de n nouveaux élèves issus de p lycées. Alors que les élèves peuvent avoir spontanément tendance à se regrouper selon leurs lycées d'origine (ce qui définit la partition \mathcal{P} dans notre modélisation), on cherche à les mélanger grâce aux activités à organiser. Les activités forment alors les classes de la partition \mathcal{Q} cherchée ; la distance $\theta(\mathcal{P}, \mathcal{Q})$ fournit une mesure du mélange obtenu, ce qui conduit à maximiser cette distance, \mathcal{P} étant donnée et le nombre d'activités (le paramètre q) pouvant de nouveau être fixé, majoré ou totalement libre.

Cet article est organisé de la façon suivante. Tout d'abord, Section 2, on donne la définition formelle de la distance de transfert ainsi que celle de la notion complémentaire, appelée *concordance*, correspondant au nombre maximum d'éléments qu'il n'est pas nécessaire de déplacer pour transformer \mathcal{P} en \mathcal{Q} ou réciproquement \mathcal{Q} en

\mathcal{P} . On rappelle ensuite comment on calcule la distance de transfert. En Section 3, on établit la concordance minimum entre une partition \mathcal{P} donnée et une partition quelconque, puis, en Section 4, la concordance minimum entre une partition \mathcal{P} donnée et une partition dont le nombre de classes est donné ou majoré par un entier q . Enfin on exprime les résultats obtenus en termes de distance en Section 5.

2. LA DISTANCE DE TRANSFERT

2.1. DÉFINITIONS ET NOTATIONS

Soit X un ensemble fini à n éléments. Rappelons qu'une partition \mathcal{P} sur X à p classes ($p \in \{1, \dots, n\}$) est constituée de p sous-ensembles P_1, P_2, \dots, P_p de X appelés classes de \mathcal{P} , qui vérifient :

$$\begin{aligned} \forall i \in \{1, \dots, p\}, P_i \neq \emptyset \\ \forall (i, j) \in \{1, \dots, p\}^2, i \neq j, P_i \cap P_j = \emptyset \\ \bigcup_{i=1}^p P_i = X. \end{aligned}$$

Soient \mathcal{P} et \mathcal{Q} deux partitions sur X à p et q classes. On note $P_i, 1 \leq i \leq p$, les classes de \mathcal{P} et $Q_j, 1 \leq j \leq q$, celles de \mathcal{Q} . On adopte comme représentation de l'ensemble X et de ces deux partitions un tableau $p \times q$ dans lequel les lignes sont les classes de \mathcal{P} et les colonnes celles de \mathcal{Q} (cf. Figure 1).

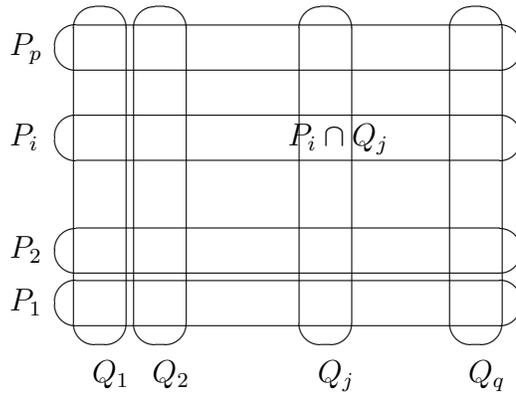


FIGURE 1. Représentation de X, \mathcal{P} et \mathcal{Q}

Soit Σ l'ensemble des couplages entre les classes de \mathcal{P} et celles de \mathcal{Q} . Un couplage σ appartenant à Σ est un ensemble de couples $(P_i, Q_j), i \in \{1, \dots, p\}$ et $j \in \{1, \dots, q\}$, tel que chaque classe apparaît au plus une fois. Si (P_i, Q_j) appartient à σ , on dit alors que σ couple P_i avec Q_j . Par souci de simplicité, on identifiera les classes et leurs indices et on dira donc, pour $i \in \{1, \dots, p\}$ et $j \in \{1, \dots, q\}$, que (i, j) appartient à σ si (P_i, Q_j) appartient à σ et réciproquement.

La σ -concordance c_σ entre \mathcal{P} et \mathcal{Q} est le nombre d'éléments classés dans des classes correspondantes suivant σ :

$$c_\sigma(\mathcal{P}, \mathcal{Q}) = \sum_{(i,j) \in \sigma} |P_i \cap Q_j|.$$

La concordance c entre \mathcal{P} et \mathcal{Q} est alors définie par :

$$c(\mathcal{P}, \mathcal{Q}) = \max_{\sigma \in \Sigma} c_\sigma(\mathcal{P}, \mathcal{Q}).$$

On calcule le nombre de transferts t_σ entre \mathcal{P} et \mathcal{Q} par :

$$t_\sigma(\mathcal{P}, \mathcal{Q}) = n - c_\sigma(\mathcal{P}, \mathcal{Q}).$$

La distance de transfert θ entre \mathcal{P} et \mathcal{Q} est alors donnée par :

$$\theta(\mathcal{P}, \mathcal{Q}) = \min_{\sigma \in \Sigma} t_\sigma(\mathcal{P}, \mathcal{Q}).$$

C'est le nombre minimum de transferts d'un élément d'une classe dans une autre pour passer d'une partition à l'autre. On a :

$$\theta(\mathcal{P}, \mathcal{Q}) = n - c(\mathcal{P}, \mathcal{Q}).$$

On peut facilement vérifier que la distance de transfert est bien une distance.

2.2. CALCUL DE LA DISTANCE DE TRANSFERT

On appellera *pseudo-partition* un système de classes disjointes dont l'union vaut X mais qui peuvent éventuellement être vides. Une pseudo-partition est donc définie comme une partition en relâchant la contrainte sur la non-vacuité des classes. On note $m = \max(p, q)$ et on considère les deux pseudo-partitions $\tilde{\mathcal{P}}$ et $\tilde{\mathcal{Q}}$ obtenues en ajoutant respectivement $m - p$ et $m - q$ classes vides P_{p+1}, \dots, P_m et Q_{q+1}, \dots, Q_m à \mathcal{P} et \mathcal{Q} (on ajoute des classes vides à la partition ayant le plus petit nombre de classes de façon à obtenir deux pseudo-partitions possédant le même nombre de classes).

On considère le graphe biparti complet $K_{m,m}$ ayant pour sommets les classes de $\tilde{\mathcal{P}}$ et de $\tilde{\mathcal{Q}}$, les classes de $\tilde{\mathcal{P}}$ formant un des ensembles de la bipartition et les classes de $\tilde{\mathcal{Q}}$ l'autre.

Soient Υ , Γ , Λ et Δ des applications de $\{1, \dots, m\}^2$ dans \mathbb{N} définies par :

$$\begin{aligned} \Upsilon(i, j) &= |P_i \cap Q_j| \\ \Gamma(i, j) &= |P_i| - |P_i \cap Q_j| \\ \Lambda(i, j) &= |Q_j| - |P_i \cap Q_j| \\ \Delta(i, j) &= |P_i| + |Q_j| - 2 |P_i \cap Q_j|. \end{aligned}$$

On rappelle qu'un *couplage* M dans un graphe $G = (V, E)$ est un sous-ensemble de E tel que deux éléments de M (i.e. deux arêtes de G) sont disjoints : $\forall \{x, y\} \in M, \forall \{z, t\} \in M, x \neq z, x \neq t, y \neq z, y \neq t$. Un couplage M est dit *parfait* si tout sommet de G est l'extrémité d'une arête de M (cf. [Lovasz, Plummer, 1986] pour des détails sur les couplages).

THÉORÈME 1. Soit σ un couplage dans $K_{m,m}$. Les propositions suivantes sont équivalentes :

- a) σ minimise le nombre de transfert $t_\sigma(\mathcal{P}, \mathcal{Q})$;
- b) σ maximise la σ -concordance $c_\sigma(\mathcal{P}, \mathcal{Q})$;
- c) σ définit un couplage parfait de poids maximum dans $K_{m,m}$ pondéré par Υ ; la valeur de ce couplage est égale à $c(\mathcal{P}, \mathcal{Q})$;
- d) σ définit un couplage parfait de poids minimum dans $K_{m,m}$ pondéré par Γ ; la valeur de ce couplage est égale à $\theta(\mathcal{P}, \mathcal{Q})$;
- e) σ définit un couplage parfait de poids minimum dans $K_{m,m}$ pondéré par Λ ; la valeur de ce couplage est égale à $\theta(\mathcal{P}, \mathcal{Q})$;
- f) σ définit un couplage parfait de poids minimum dans $K_{m,m}$ pondéré par Δ ; la valeur de ce couplage est égale à $2\theta(\mathcal{P}, \mathcal{Q})$.

Proof. (Théorème 1). Tout d'abord il est clair qu'un couplage qui maximise $c_\sigma(\mathcal{P}, \mathcal{Q})$ couple un maximum de classes et donc correspond à un couplage parfait dans $K_{m,m}$. L'équivalence entre a) et c) a été prouvée par Day [1981]. Ce théorème découle simplement des relations suivantes :

$$\begin{aligned} \theta(\mathcal{P}, \mathcal{Q}) &= n - c(\mathcal{P}, \mathcal{Q}) = \min_{\sigma \in \Sigma} (n - \sum_{(i,j) \in \sigma} \Upsilon(i,j)) = \min_{\sigma \in \Sigma} \sum_{(i,j) \in \sigma} \Gamma(i,j) \\ &= \min_{\sigma \in \Sigma} \sum_{(i,j) \in \sigma} \Lambda(i,j) = \min_{\sigma \in \Sigma} \frac{1}{2} \sum_{(i,j) \in \sigma} \Delta(i,j). \end{aligned}$$

□

Le calcul de $\theta(\mathcal{P}, \mathcal{Q})$ est donc un problème d'affectation (des classes de $\tilde{\mathcal{P}}$ aux classes de $\tilde{\mathcal{Q}}$) qui consiste à chercher un couplage parfait (à m arêtes) de poids minimum dans $K_{m,m}$ pondéré par Γ (ou de manière équivalente par Λ ou Δ) ou un couplage de poids maximum dans $K_{m,m}$ pondéré par Υ . La construction du graphe biparti est en $O(n^2)$. Le problème du couplage de poids minimum dans un graphe biparti complet est bien connu en recherche opérationnelle. Il peut être résolu par la *méthode hongroise* [Kuhn, 1955, 1956], un algorithme de complexité polynomiale en $O(m^3)$. Un lecteur intéressé pourra trouver des détails dans [Ahuja, Magnanti, Orlin, 1993].

Exemple 1. Soient $\mathcal{P} = (1, 2, 3|4, 5, 6|7, 8, 9)$ et $\mathcal{Q} = (1, 3, 5, 6|2, 7, 9|4|8)$ deux partitions sur $X = \{1, 2, \dots, 8\}$. Les deux tableaux suivants correspondent respectivement aux intersections et aux différences symétriques entre les classes de \mathcal{P} et celles de \mathcal{Q} . Un couplage optimal figure en gras. Il donne $\theta(\mathcal{P}, \mathcal{Q}) = 4$.

Υ	1,3,5,6	2,7,9	4	8	Δ	1,3,5,6	2,7,9	4	8
1,2,3	2	1	0	0	1,2,3	3	4	4	4
4,5,6	2	0	1	0	4,5,6	3	6	2	4
7,8,9	0	2	0	1	7,8,9	7	2	4	2
\emptyset	0	0	0	0	\emptyset	4	3	1	1

À ce couplage maximum correspond par exemple la série de transferts suivante :

$$(1, \mathbf{2}, 3|4, 5, 6|7, 8, 9) \rightarrow (1, 3|4, \mathbf{5}, 6|2, 7, 8, 9) \rightarrow (1, 3, 5|4, \mathbf{6}|2, 7, 8, 9) \\ \rightarrow (1, 3, 5, 6|4|2, 7, \mathbf{8}, 9) \rightarrow (1, 3, 5, 6|4|2, 7, 9|8)$$

3. CONCORDANCE MINIMUM AVEC UNE PARTITION \mathcal{P} DONNÉE

Dans cette partie, on cherche à déterminer la concordance minimum $c_{\min}(\mathcal{P})$ pour une partition \mathcal{P} donnée :

$$c_{\min}(\mathcal{P}) = \min_{\mathcal{Q}} c(\mathcal{P}, \mathcal{Q}).$$

On suppose que les classes de \mathcal{P} sont rangées dans l'ordre des cardinaux croissants ($n_1 \leq n_2 \leq \dots \leq n_p$). Soit i_0 l'indice vérifiant $n_{i_0} - i_0 = \min_{1 \leq i \leq p} (n_i - i)$. Soient \mathcal{Q}^1 et \mathcal{Q}^2 deux partitions définies comme suit : la classe Q_1^1 contient tous les éléments des classes P_i , $1 \leq i \leq i_0$, toutes les autres classes de \mathcal{Q}^1 sont des singletons ; \mathcal{Q}^2 est la partition à n classes. On se propose de démontrer le théorème suivant :

THÉORÈME 2. *Soit \mathcal{P} une partition à p classes possédant respectivement n_1, n_2, \dots, n_p éléments avec $n_1 \leq n_2 \leq \dots \leq n_p$. Alors*

$$c_{\min}(\mathcal{P}) = \min[0, \min_{1 \leq i \leq p} (n_i - i)] + p,$$

et cette valeur est atteinte par \mathcal{Q}^1 ou \mathcal{Q}^2 .

Illustrons ce résultat par un exemple.

Exemple 2. Soit $X = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$, et la partition à 4 classes \mathcal{P} décrite dans le tableau suivant :

i	1	2	3	4
P_i	1	2, 3	4, 5	6, 7, 8, 9
n_i	1	2	2	4
$n_i - i$	0	0	-1	0

L'indice i_0 vaut 3. On construit donc les partitions \mathcal{Q}^1 et \mathcal{Q}^2 de la façon suivante. On regroupe les éléments des 3 premières classes de \mathcal{P} dans la première classe de \mathcal{Q}^1 , et on disperse les autres : $\mathcal{Q}^1 = (12345|6|7|8|9)$. La partition \mathcal{Q}^2 est la partition à 9 singletons $(1|2|3|4|5|6|7|8|9)$. Calculons maintenant les concordances entre \mathcal{P} et ces deux partitions.

Un meilleur couplage entre les classes de \mathcal{P} et celles de \mathcal{Q}^1 associe Q_1^1 avec la classe P_{i_0} (P_{i_0} est la classe la plus grosse contenue dans Q_1^1 , puisque les classes sont ordonnées par cardinaux croissants). Ensuite, on pourra au mieux coupler chaque P_i , $i_0 < i \leq p$, avec une classe Q_j^1 , $1 < j \leq q$, telle que $|P_i \cap Q_j^1| = 1$. Ainsi, la concordance entre \mathcal{P} et \mathcal{Q}^1 vaut $n_{i_0} + (p - i_0)$. Dans notre exemple, elle consiste à coupler P_3 avec Q_1^1 et P_4 avec une des autres classes de \mathcal{Q}^1 et vaut donc 3.

Un meilleur couplage entre les classes de \mathcal{P} et celles de \mathcal{Q}^2 sera construit en couplant chaque P_i , $1 \leq i \leq p$, avec une classe Q_j^2 , $1 \leq j \leq q$, telle que $|P_i \cap Q_j^2| = 1$. On obtient donc une concordance égale à p , soit dans notre exemple $c(\mathcal{P}, \mathcal{Q}^2) = p = 4$ (on associe par exemple P_1 avec Q_1^2 , P_2 avec Q_2^2 , P_3 avec Q_4^2 et P_4 avec Q_6^2).

La concordance minimum à la partition \mathcal{P} est donc atteinte par \mathcal{Q}^1 et vaut 3.

Avant de commencer la démonstration de ce théorème, commençons par énoncer et prouver le lemme suivant.

On considère une partition \mathcal{P} sur X et un élément $a \in X$. Soit k l'indice de la classe de \mathcal{P} contenant l'élément a . On note alors $\mathcal{P}_{\ominus a}$ la partition sur $X \setminus \{a\}$ définie par :

- si $P_k = \{a\}$, $\mathcal{P}_{\ominus a}$ est constituée des classes P_i , $i \neq k$.
- sinon $\mathcal{P}_{\ominus a}$ est constituée des classes P_i , $i \neq k$, ainsi que de la classe $P_k \setminus \{a\}$.

LEMME 1. *On considère une partition \mathcal{P} sur X et un élément $a \in X$. Alors :*

$$c_{\min}(\mathcal{P}) \geq c_{\min}(\mathcal{P}_{\ominus a}).$$

Proof. (Lemme 1). Soit \mathcal{Q} une partition sur X et $\mathcal{Q}_{\ominus a}$ définie à partir de \mathcal{Q} comme décrit plus haut. Soit $\hat{\sigma}$ le couplage réalisant la concordance maximum entre $\mathcal{P}_{\ominus a}$ et $\mathcal{Q}_{\ominus a}$. Par définition, on a

$$c_{\min}(\mathcal{P}_{\ominus a}) \leq c(\mathcal{P}_{\ominus a}, \mathcal{Q}_{\ominus a}) = c_{\hat{\sigma}}(\mathcal{P}_{\ominus a}, \mathcal{Q}_{\ominus a})$$

On note encore $\hat{\sigma}$ le couplage entre \mathcal{P} et \mathcal{Q} obtenu à partir de $\hat{\sigma}$ en remettant a à sa place dans chacune des partitions ; si $P_k = \{a\}$, P_k est couplée avec une classe de \mathcal{Q} non couplée ou, s'il n'y en a pas, P_k n'est pas couplée. Alors par définition de la concordance :

$$c(\mathcal{P}, \mathcal{Q}) = \max_{\sigma} c_{\sigma}(\mathcal{P}, \mathcal{Q}) \geq c_{\hat{\sigma}}(\mathcal{P}, \mathcal{Q}).$$

Or on remarque l'inégalité

$$c_{\hat{\sigma}}(\mathcal{P}, \mathcal{Q}) \geq c_{\hat{\sigma}}(\mathcal{P}_{\ominus a}, \mathcal{Q}_{\ominus a})$$

puisque soit $\hat{\sigma}$ couple les classes de \mathcal{P} et \mathcal{Q} contenant a et alors

$$c_{\hat{\sigma}}(\mathcal{P}, \mathcal{Q}) = c_{\hat{\sigma}}(\mathcal{P}_{\ominus a}, \mathcal{Q}_{\ominus a}) + 1,$$

soit ce n'est pas le cas et alors

$$c_{\hat{\sigma}}(\mathcal{P}, \mathcal{Q}) = c_{\hat{\sigma}}(\mathcal{P}_{\ominus a}, \mathcal{Q}_{\ominus a}).$$

Finalement on a :

$$c(\mathcal{P}, \mathcal{Q}) \geq c_{\hat{\sigma}}(\mathcal{P}, \mathcal{Q}) \geq c_{\hat{\sigma}}(\mathcal{P}_{\ominus a}, \mathcal{Q}_{\ominus a}) \geq c_{\min}(\mathcal{P}_{\ominus a}),$$

et ce, quelle que soit la partition \mathcal{Q} considérée, ce qui implique

$$c_{\min}(\mathcal{P}) = \min_{\mathcal{Q}} c(\mathcal{P}, \mathcal{Q}) \geq c_{\min}(\mathcal{P}_{\ominus a}).$$

□

Démontrons maintenant le Théorème 2.

Proof. (Théorème 2). On note $f(\mathcal{P}) = \min[0, \min_{1 \leq i \leq p}(n_i - i)] + p$.

Tout d'abord, montrons que $\forall \mathcal{P}, f(\mathcal{P}) \geq c_{\min}(\mathcal{P})$.

La concordance entre \mathcal{P} et \mathcal{Q}^1 , partition décrite plus haut, vaut $c(\mathcal{P}, \mathcal{Q}^1) = p - i_0 + n_{i_0}$ (les couplages optimaux entre les classes de \mathcal{P} et celles de \mathcal{Q}^1 couplent P_{i_0} avec Q_1^1 , et $P_i, i > i_0$, avec une classe Q_j^1 telle que $|P_i \cap Q_j^1| = 1$). D'autre part il est immédiat que $c(\mathcal{P}, \mathcal{Q}^2) = p$. Ainsi

$$c_{\min}(\mathcal{P}) = \min_{\mathcal{Q}} c(\mathcal{P}, \mathcal{Q}) \leq \min[c(\mathcal{P}, \mathcal{Q}^1), c(\mathcal{P}, \mathcal{Q}^2)] = \min[p, p + \min_{1 \leq i \leq p}(n_i - i)] = f(\mathcal{P}).$$

Raisonnons maintenant par récurrence sur n pour montrer que $c_{\min}(\mathcal{P}) \geq f(\mathcal{P})$.

• Initialisation : si $n = 1$, l'ensemble des partitions sur X est réduit à la partition à une classe à un élément et donc il est clair que $c_{\min}(\mathcal{P}) = 1$. De plus puisque $p = 1$ et $n_1 = 1$, on a

$$f(\mathcal{P}) = \min[0, \min_{1 \leq i \leq p}(n_i - i)] + p = 1$$

et donc $c_{\min}(\mathcal{P}) = f(\mathcal{P})$.

• Hérité : on suppose maintenant que le théorème est vrai pour toute partition sur au plus $n - 1$ éléments. On considère une partition \mathcal{P} quelconque sur n éléments.

- **cas 1** : $n_1 \geq 2$

On considère un élément a de la classe P_1 et la partition $\mathcal{P}' = \mathcal{P} \ominus a$. La partition \mathcal{P}' possède toujours p classes notées $P'_i, i \in \{1, \dots, p\}$, de cardinaux respectifs n'_i avec $n'_1 \leq n'_2 \leq \dots \leq n'_p$. D'après l'hypothèse de récurrence,

$$c_{\min}(\mathcal{P}') = \min[0, \min_{1 \leq i \leq p}(n'_i - i)] + p = \min[0, \min_{2 \leq i \leq p}(n_i - i), n_1 - 2] + p.$$

Puisque $n_1 - 2 \geq 0$, on a donc

$$c_{\min}(\mathcal{P}') = f(\mathcal{P}).$$

Or, d'après le Lemme 1, on sait que $c_{\min}(\mathcal{P}) \geq c_{\min}(\mathcal{P}')$ et donc on obtient finalement

$$c_{\min}(\mathcal{P}) \geq f(\mathcal{P}).$$

- **cas 2** : $n_1 = 1$

- **cas 2.1** : $\min_{i \in \{1, \dots, p\}}(n_i - i) < 0$

On note à nouveau i_0 l'indice vérifiant $n_{i_0} - i_0 = \min_{i \in \{1, \dots, p\}}(n_i - i)$. Soient a l'unique élément de P_1 et $\mathcal{P}' = \mathcal{P} \ominus a$. La partition \mathcal{P}' possède $p - 1$ classes de cardinaux respectifs $n'_i = n_{i+1}, i \in \{1, \dots, p - 1\}$. On a donc

$$c_{\min}(\mathcal{P}') = \min[0, \min_{1 \leq i \leq p-1}(n_{i+1} - i)] + p - 1.$$

Or, puisque $i_0 \neq 1$,

$$\min_{1 \leq i \leq p-1} (n_{i+1} - i) = \min_{2 \leq i \leq p} (n_i - i) + 1 = n_{i_0} - i_0 + 1$$

et, d'après l'hypothèse, $n_{i_0} - i_0 < 0$. Finalement,

$$c_{\min}(\mathcal{P}') = n_{i_0} - i_0 + 1 + p - 1 = n_{i_0} - i_0 + p = f(\mathcal{P})$$

et on conclut comme pour le cas 1.

- **cas 2.2 :** $\forall i \in \{1, \dots, p\}$, $n_i - i \geq 0$ et $\exists i$ tel que $n_i - i > 0$

Soit i_0 le plus petit indice vérifiant $n_i - i > 0$. Soient $a \in P_{i_0}$ et $\mathcal{P}' = \mathcal{P} \ominus a$. La partition \mathcal{P}' possède p classes de cardinaux n'_i définis comme suit : si $i \in \{1, \dots, p\}$, $i \neq i_0$, $n'_i = n_i$ et $n'_{i_0} = n_{i_0} - 1$. Remarquons que la suite des n'_i reste ordonnée dans l'ordre croissant ; en effet d'après la définition de i_0 , on a $n_{i_0} - i_0 > 0$ et $n_{i_0-1} - (i_0 - 1) \leq 0$, ce qui implique que $n_{i_0} > n_{i_0-1}$. Alors d'après l'hypothèse de récurrence :

$$\begin{aligned} c_{\min}(\mathcal{P}') &= \min[0, \min_{1 \leq i \leq p} (n'_i - i)] + p \\ \Leftrightarrow c_{\min}(\mathcal{P}') &= \min[0, \min_{1 \leq i \leq p, i \neq i_0} (n_i - i), n_{i_0} - 1 - i_0] + p \end{aligned}$$

et donc

$$c_{\min}(\mathcal{P}') = p = f(\mathcal{P}).$$

On conclut à nouveau comme pour le cas 2.1.

- **cas 2.3 :** $\forall i \in \{1, \dots, p\}$, $n_i - i = 0$

Soit a l'élément unique de la classe P_1 . On considère une partition \mathcal{Q} quelconque sur n éléments. Appelons Q_1 la classe de \mathcal{Q} qui contient l'élément a . On construit deux partitions \mathcal{P}' et \mathcal{Q}' à partir des partitions \mathcal{P} et \mathcal{Q} en enlevant un élément de chaque ensemble $Q_1 \cap P_i$ pour tout $i \in \{1, \dots, p\}$ lorsque c'est possible. La partition \mathcal{P}' possède alors $p - 1$ classes (notées P'_i , $1 \leq i \leq p - 1$) de cardinaux respectifs $n'_i = n_{i+1}$ ou $n'_i = n_{i+1} - 1$. On remarque que $n'_1 \leq n'_2 \leq \dots \leq n'_{p-1}$ puisque dans ce cas pour tout $i \in \{1, \dots, p\}$, $n_i > n_{i-1}$. La partition \mathcal{Q}' est identique à la partition \mathcal{Q} excepté pour la première classe, de laquelle ont été retirés un certain nombre d'éléments (dont a).

On se propose de montrer que $c(\mathcal{P}', \mathcal{Q}') \geq p - 1$. Puisque n'_i vaut n_{i+1} ou $n_{i+1} - 1$ pour tout $i \in \{1, \dots, p - 1\}$ et qu'on a, pour tout $i \in \{1, \dots, p\}$, $n_i = i$ alors, pour tout $i \in \{1, \dots, p - 1\}$, n'_i vaut soit i , soit $i + 1$, et donc $n'_i - i \in \{0, 1\}$. D'après l'hypothèse de récurrence,

$$c_{\min}(\mathcal{P}') = \min[0, \min_{1 \leq i \leq p-1} (n'_i - i)] + p - 1 = p - 1.$$

Ainsi, par définition de $c_{\min}(\mathcal{P}')$, on obtient

$$c(\mathcal{P}', \mathcal{Q}') \geq p - 1.$$

Il existe donc un couplage σ entre les classes de \mathcal{P}' et celles de \mathcal{Q}' tel que $c_\sigma(\mathcal{P}', \mathcal{Q}') \geq p - 1$. Par définition,

$$c_\sigma(\mathcal{P}', \mathcal{Q}') = \sum_{(i,j) \in \sigma} |P'_i \cap Q'_j|.$$

On observe, pour tout couple (i, j) , $i \neq 1$, et $j \neq 1$, que $P'_i \cap Q'_j = P_i \cap Q_j$. Soit P'_{i_0} , $i_0 \geq 2$, la classe couplée avec Q'_1 lorsque celle-ci est couplée. On distingue les deux cas suivants :

- $P'_{i_0} \cap Q'_1 \neq \emptyset$

Alors, d'après la construction de \mathcal{P}' et \mathcal{Q}' ,

$$|P'_{i_0} \cap Q'_1| = |P_{i_0} \cap Q_1| - 1.$$

Finalemnt, en prolongeant σ en ne couplant pas P_1 si toutes les classes de \mathcal{Q} sont déjà couplées ou avec une des classes restantes sinon, on obtient :

$$c_\sigma(\mathcal{P}, \mathcal{Q}) = \sum_{(i,j) \in \sigma} |P_i \cap Q_j| = c_\sigma(\mathcal{P}', \mathcal{Q}') + 1.$$

- $P'_{i_0} \cap Q'_1 = \emptyset$ ou Q'_1 n'est pas couplée

On prolonge σ en couplant P_1 à Q_1 et on obtient de la même façon

$$c_\sigma(\mathcal{P}, \mathcal{Q}) = c_\sigma(\mathcal{P}', \mathcal{Q}') + 1.$$

Donc, pour ces deux cas,

$$c(\mathcal{P}, \mathcal{Q}) \geq c_\sigma(\mathcal{P}, \mathcal{Q}) = c_\sigma(\mathcal{P}', \mathcal{Q}') + 1 \geq p - 1 + 1 = p$$

Ainsi on a montré, pour \mathcal{Q} quelconque, l'inégalité $c(\mathcal{P}, \mathcal{Q}) \geq p$, et on obtient à nouveau dans le cas 2.2.3 l'inégalité :

$$c_{\min}(\mathcal{P}) \geq f(\mathcal{P}).$$

Finalemnt, pour toute partition \mathcal{P} sur n éléments, on a montré :

$$c_{\min}(\mathcal{P}) \geq f(\mathcal{P}).$$

D'après la première partie de la démonstration, on sait que $c_{\min}(\mathcal{P}) \leq f(\mathcal{P})$, donc il vient :

$$\forall \mathcal{P} \text{ sur } n \text{ éléments, } c_{\min}(\mathcal{P}) = f(\mathcal{P}),$$

ce qui achève la démonstration par récurrence. \square

Exemple 3. On donne ici, pour toute partition \mathcal{P} sur $X = \{1, 2, 3, 4, 5, 6\}$, les partitions \mathcal{Q} réalisant $c_{\min}(\mathcal{P})$. Étant donné que seuls les cardinaux des classes de \mathcal{P} entrent en compte dans le calcul de $c_{\min}(\mathcal{P})$, on note les éléments par l'indice de leur classe dans \mathcal{P} , la permutation des éléments d'une même classe ne changeant rien au calcul de la concordance.

On remarque que \mathcal{Q}^1 ou \mathcal{Q}^2 apparaît systématiquement dans ce tableau, comme l'établit le théorème 2. Il n'existe pas cependant de stratégie générale pour savoir laquelle de ces deux partitions sera à concordance minimum par rapport à la partition de départ. On remarque de plus que d'autres partitions peuvent correspondre à ce minimum, et que le nombre de ces partitions varie beaucoup suivant \mathcal{P} .

\mathcal{P}	$c_{\min}(\mathcal{P})$	$\mathcal{Q} \mid c(\mathcal{P}, \mathcal{Q}) = c_{\min}(\mathcal{P})$
(1 2 3 4 5 6)	1	(123456)
(1 2 3 4 55)	2	(123455), (12345 5), (2345 15), (1345 25), (1245 35), (1235 45), (125 345), (135 245), (145 235), (1234 5 5)
(1 2 33 44)	2	(123344), (1234 34), (134 234)
(11 22 33)	2	(112233), (123 123)
(1 2 3 444)	2	(12344 4), (1234 4 4)
(1 22 333)	3	(122333), (1233 23), (123 23 3) (13 23 23), (1 2 2 3 3 3)
(111 222)	2	(12 12 12), (12 12 1 2), (12 1 1 2 2), (1 1 1 2 2 2)
(1 2 3333)	2	(1233 3 3), (123 3 3 3)
(11 2222)	2	(12 12 2 2), (1 1 2 2 2 2)
(1 22222)	2	(122 2 2 2), (12 2 2 2 2)
(111111)	1	(1 1 1 1 1 1)

4. CONCORDANCE MINIMUM ENTRE UNE PARTITION \mathcal{P} DONNÉE ET UNE PARTITION À AU PLUS q CLASSES

Dans cette partie, on cherche à déterminer la concordance minimum entre une partition \mathcal{P} donnée et une partition \mathcal{Q} à au plus q classes, q étant donné. On verra que dans le cas où q est plus petit que le cardinal de la plus grosse classe de \mathcal{P} , cette concordance est égale à la concordance minimum entre une partition \mathcal{P} donnée et une partition \mathcal{Q} à exactement q classes. On note

$$c_{\min}(\mathcal{P}, \leq q) = \min\{c(\mathcal{P}, \mathcal{Q}) : \mathcal{Q} \text{ partition à au plus } q \text{ classes}\}$$

et

$$c_{\min}(\mathcal{P}, q) = \min\{c(\mathcal{P}, \mathcal{Q}) : \mathcal{Q} \text{ partition à } q \text{ classes}\}$$

On rappelle qu'une pseudo-partition est définie comme une partition en relâchant la contrainte sur la non-vacuité des classes. Ainsi toute pseudo-partition à q classes correspond à une partition à au plus q classes obtenue en supprimant les classes vides de la pseudo-partition, et réciproquement toute partition à au plus q classes correspond à une pseudo-partition à q classes obtenue en ajoutant un certain nombre de classes vides à la partition. On étend alors la définition de la concordance aux pseudo-partitions, et, puisque les classes vides n'entrent pas en compte dans le calcul de la concordance, on a pour toute partition \mathcal{P} et toute pseudo-partition $\tilde{\mathcal{Q}}$:

$$c(\mathcal{P}, \tilde{\mathcal{Q}}) = c(\mathcal{P}, \mathcal{Q})$$

où \mathcal{Q} est la partition correspondant à $\tilde{\mathcal{Q}}$. On a donc

$$\begin{aligned} c_{\min}(\mathcal{P}, \leq q) &= \min\{c(\mathcal{P}, \mathcal{Q}) : \mathcal{Q} \text{ partition à au plus } q \text{ classes}\} \\ &= \min\{c(\mathcal{P}, \tilde{\mathcal{Q}}) : \tilde{\mathcal{Q}} \text{ pseudo-partition à } q \text{ classes}\} \end{aligned}$$

On note à nouveau p le nombre de classes de \mathcal{P} , P_1, P_2, \dots, P_p ses classes et n_1, n_2, \dots, n_p les cardinaux de celles-ci. On pose de plus pour tout $i \in \{1, \dots, p\}$, $n_i = t_i q + r_i$ avec $r_i \in \{0, \dots, q-1\}$.

Soit $\hat{\mathcal{Q}}$ la pseudo-partition sur X à q classes (représentée Figure 2) construite à partir de \mathcal{P} en distribuant un par un les éléments de chaque classe de \mathcal{P} dans les q classes de $\hat{\mathcal{Q}}$ en commençant toujours par la classe $\hat{\mathcal{Q}}_1$. Chaque classe $\hat{\mathcal{Q}}_j$ contient $t_i + \epsilon_{i,j}$ éléments appartenant à la classe P_i , $\forall i \in \{1, \dots, p\}$, avec $\epsilon_{i,j} = 1$ si $r_i \geq j$ et $\epsilon_{i,j} = 0$ si $r_i < j$. En effet on peut distribuer t_i éléments de chaque classe P_i dans chacune des classes de $\hat{\mathcal{Q}}$; restent ensuite r_i éléments qu'on répartit dans les r_i premières classes de $\hat{\mathcal{Q}}$.

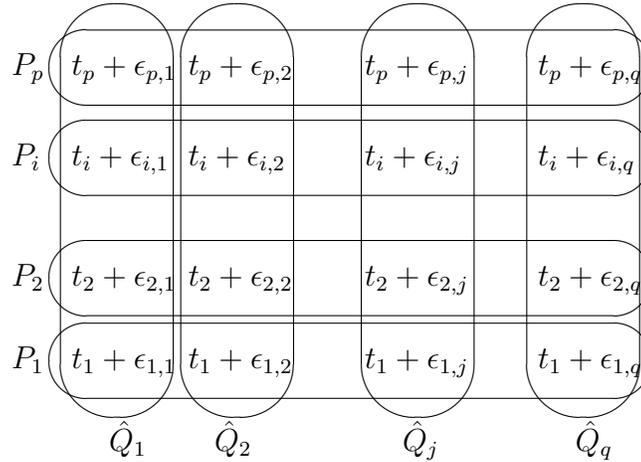


FIGURE 2. Partitions \mathcal{P} et $\hat{\mathcal{Q}}$

On distingue deux cas suivant les valeurs respectives de p et q .

4.1. CAS 1. VALEUR DE $c_{\min}(\mathcal{P}, \leq q)$ QUAND $p \leq q$

On se propose de démontrer le théorème suivant, établissant la valeur minimum de la concordance entre une partition \mathcal{P} donnée à p classes et une partitions à au plus q classes dans le cas où $p \leq q$.

THÉORÈME 3. *Soit \mathcal{P} une partition à p classes sur X et q un entier vérifiant $p \leq q$. On note n_i , $i \in \{1, \dots, p\}$, les cardinaux des classes de \mathcal{P} et r_i , $i \in \{1, \dots, p\}$, les restes respectifs de la division entière des n_i par q et on suppose que les classes sont indicées dans l'ordre croissant des r_i ($r_1 \leq r_2 \leq \dots \leq r_p$). Alors*

$$c_{\min}(\mathcal{P}, \leq q) = \sum_{i=1}^p \left\lfloor \frac{n_i}{q} \right\rfloor + \min(0, \min_{1 \leq i \leq p} (r_i - i)) + p.$$

On rappelle que $t_i = \left\lfloor \frac{n_i}{q} \right\rfloor$ pour tout $i \in \{1, \dots, p\}$ et on pose

$$f(\mathcal{P}, q) = \sum_{i=1}^p t_i + \min(0, \min_{1 \leq i \leq p} (r_i - i)) + p.$$

Afin de démontrer ce théorème, nous aurons besoin de certains résultats préliminaires.

PROPOSITION 1. *Pour toute partition \mathcal{P} à p classes et tout entier q tel que $p \leq q$ et pour la partition $\hat{\mathcal{Q}}$ définie à partir de \mathcal{P} comme décrit plus haut, on a :*

$$c(\mathcal{P}, \hat{\mathcal{Q}}) = f(\mathcal{P}, q).$$

Proof. (Proposition 1). Montrons tout d'abord que $c(\mathcal{P}, \hat{\mathcal{Q}}) \leq f(\mathcal{P}, q)$. On considère un couplage quelconque σ entre les classes de \mathcal{P} et celles de $\hat{\mathcal{Q}}$. Montrons par récurrence sur les classes de \mathcal{P} que $c_\sigma(\mathcal{P}, \hat{\mathcal{Q}}) \leq f(\mathcal{P}, q)$. Pour $1 \leq k \leq p$, on pose $m_k = \min(0, \min_{1 \leq i \leq k} (r_i - i))$ ($m_k \leq 0$) et on note σ_k le couplage σ restreint aux k premières classes de \mathcal{P} ($(i, j) \in \sigma_k$ si $(i, j) \in \sigma$ et $i \leq k$). Montrons que $\forall k \in \{1, \dots, p\}$, $\sum_{(i,j) \in \sigma_k} |P_i \cap \hat{Q}_j| \leq \sum_{i=1}^k t_i + k + m_k$.

• Initialisation : montrons que l'inégalité est vérifiée pour $k = 1$: on note ℓ l'indice de la classe couplée à la classe P_1 par σ . Alors

$$\sum_{(i,j) \in \sigma_1} |P_i \cap \hat{Q}_j| = |P_1 \cap \hat{Q}_\ell| = t_1 + \epsilon_{1,\ell}$$

avec $\epsilon_{1,\ell} = 0$ si $r_1 < \ell$ et $\epsilon_{1,\ell} = 1$ sinon. Or $m_1 = \min(0, r_1 - 1) = -1$ si $r_1 = 0$ et $m_1 = 0$ sinon, donc :

- si $r_1 = 0$, $|P_1 \cap \hat{Q}_\ell| = t_1 = t_1 + 1 + m_1$
- si $r_1 \neq 0$, $|P_1 \cap \hat{Q}_\ell| \leq t_1 + 1 = t_1 + 1 + m_1$.

Ainsi dans tous les cas

$$|P_1 \cap \hat{Q}_\ell| \leq t_1 + 1 + m_1.$$

• Hérité : on suppose que $\sum_{(i,j) \in \sigma_{k-1}} |P_i \cap \hat{Q}_j| \leq \sum_{i=1}^{k-1} t_i + k - 1 + m_{k-1}$. Montrons que l'inégalité est alors vérifiée à l'ordre k ; ℓ désignera maintenant l'indice de la classe couplée à P_k par σ s'il y en a une, ou une classe non couplée de $\hat{\mathcal{Q}}$ sinon.

- **cas 1** : $\sum_{(i,j) \in \sigma_{k-1}} |P_i \cap \hat{Q}_j| \leq \sum_{i=1}^{k-1} t_i + k - 1 + m_{k-1} - 1$

On a toujours, d'après la construction de $\hat{\mathcal{Q}}$,

$$|P_k \cap \hat{Q}_\ell| \leq t_k + 1.$$

D'autre part on a :

$$m_k = \min(0, \min_{1 \leq i \leq k} (r_i - i)) = \min(0, \min_{1 \leq i \leq k-1} (r_i - i), r_k - k) = \min(m_{k-1}, r_k - k),$$

ce qui implique l'inégalité suivante :

$$m_k \geq m_{k-1} - 1.$$

En effet : si $r_k - k < m_{k-1}$, on a alors $r_k - k = m_{k-1} - 1$ car la suite des $(r_i - i)$ décroît au maximum de 1 en 1 puisque $r_1 \leq r_2 \leq \dots \leq r_p$, et donc $m_k = m_{k-1} - 1$; sinon $m_k = m_{k-1}$. Ainsi

$$\sum_{(i,j) \in \sigma_k} |P_i \cap \hat{Q}_j| = \left(\sum_{(i,j) \in \sigma_{k-1}} |P_i \cap \hat{Q}_j| \right) + |P_k \cap \hat{Q}_\ell|$$

$$\begin{aligned} \Rightarrow \sum_{(i,j) \in \sigma_k} |P_i \cap \hat{Q}_j| &\leq \sum_{i=1}^{k-1} t_i + k - 1 + m_{k-1} - 1 + t_k + 1 \\ &\Rightarrow \sum_{(i,j) \in \sigma_k} |P_i \cap \hat{Q}_j| \leq \sum_{i=1}^k t_i + k + m_k. \end{aligned}$$

- **cas 2** : $\sum_{(i,j) \in \sigma_{k-1}} |P_i \cap \hat{Q}_j| = \sum_{i=1}^{k-1} t_i + k - 1 + m_{k-1}$

D'après la construction de \hat{Q} ,

$$\sum_{(i,j) \in \sigma_{k-1}} |P_i \cap \hat{Q}_j| = \sum_{i=1}^{k-1} t_i + \sum_{(i,j) \in \sigma_{k-1}} \epsilon_{i,j}$$

avec $\epsilon_{i,j} = 0$ si $r_i < j$ et $\epsilon_{i,j} = 1$ sinon. Dans ce cas on a donc

$$k - 1 + m_{k-1} = \sum_{(i,j) \in \sigma_{k-1}} \epsilon_{i,j} = |\{(i,j) \in \sigma_{k-1}, j \leq r_i\}|.$$

De plus on peut remarquer que $|\{(i,j) \in \sigma_{k-1}, j \leq r_i\}| \leq r_k$ puisque tous les éléments j considérés sont distincts et inférieurs à r_k (les r_i sont ordonnés dans l'ordre croissant). On distingue alors à nouveau deux cas :

- si $k - 1 + m_{k-1} = r_k$, c'est-à-dire si $|\{(i,j) \in \sigma_{k-1}, j \leq r_i\}| = r_k$, toutes les classes \hat{Q}_j , $j \in \{1, \dots, r_k\}$, sont couplées avec une classe P_i , $i \leq k - 1$. Donc l'indice ℓ de la classe couplée à P_k par σ vérifie $\ell > r_k$ et donc $\epsilon_{k,\ell} = 0$ et $|P_k \cap \hat{Q}_\ell| = t_k$. Ainsi :

$$\sum_{(i,j) \in \sigma_k} |P_i \cap \hat{Q}_j| = \sum_{i=1}^k t_i + k - 1 + m_{k-1}.$$

Par ailleurs on sait que $m_k \geq m_{k-1} - 1$ et donc :

$$\sum_{(i,j) \in \sigma_k} |P_i \cap \hat{Q}_j| \leq \sum_{i=1}^k t_i + k + m_k.$$

- si $k - 1 + m_{k-1} < r_k$ alors $r_k - k \geq m_{k-1}$ et donc

$$m_k = \min(m_{k-1}, r_k - k) = m_{k-1}.$$

Par ailleurs on a toujours $|P_k \cap \hat{Q}_\ell| \leq t_k + 1$ et donc

$$\sum_{(i,j) \in \sigma_k} |P_i \cap \hat{Q}_j| \leq \sum_{i=1}^k t_i + 1 + k - 1 + m_k = \sum_{i=1}^k t_i + k + m_k.$$

On a ainsi démontré dans tous les cas que

$$\sum_{(i,j) \in \sigma_k} |P_i \cap \hat{Q}_j| \leq \sum_{i=1}^k t_i + k + m_k.$$

On en conclut que cette inégalité est vraie pour tout $k \in \{1, \dots, p\}$ et donc

$$c_\sigma(\mathcal{P}, \hat{\mathcal{Q}}) = \sum_{(i,j) \in \sigma} |P_i \cap \hat{Q}_j| \leq \sum_{i=1}^p t_i + k + m_p = f(\mathcal{P}, q),$$

et ce quel que soit le couplage σ considéré. Ainsi

$$c(\mathcal{P}, \hat{\mathcal{Q}}) = \max_{\sigma} c_\sigma(\mathcal{P}, \hat{\mathcal{Q}}) \leq f(\mathcal{P}, q).$$

Montrons maintenant que cette valeur peut être atteinte par un certain couplage $\hat{\sigma}$ entre les classes de \mathcal{P} et celles de $\hat{\mathcal{Q}}$. On propose l'algorithme de couplage suivant : pour i allant de 1 à p , on couple itérativement P_i à une classe \hat{Q}_j telle que $j \leq r_i$ si c'est possible et, si ce n'est pas possible, on la couple à la fin avec n'importe quelle classe de $\hat{\mathcal{Q}}$ restante. Soit $\hat{\sigma}$ un couplage obtenu par ce procédé. Montrons que $c_{\hat{\sigma}}(\mathcal{P}, \hat{\mathcal{Q}}) = f(\mathcal{P}, q)$. On procède à nouveau par récurrence sur les classes de \mathcal{P} en considérant les couplages $\hat{\sigma}_k$, $k \in \{1, \dots, p\}$, restrictions de $\hat{\sigma}$ aux k premières classes de \mathcal{P} . On se propose alors de montrer que $\forall k \in \{1, \dots, p\}$, $\sum_{(i,j) \in \hat{\sigma}_k} |P_i \cap \hat{Q}_j| = \sum_{i=1}^k t_i + k + m_k$.

- Initialisation : si $k = 1$, on appelle ℓ l'indice de la classe couplée avec P_1 par $\hat{\sigma}$. Par construction de $\hat{\mathcal{Q}}$, $|P_1 \cap \hat{Q}_\ell| = t_1 + \epsilon_{1,\ell}$ avec $\epsilon_{1,\ell} = 0$ si $r_1 < \ell$ et $\epsilon_{1,\ell} = 1$ sinon. Donc si $r_1 = 0$, $\epsilon_{1,\ell} = 0$, et si $r_1 > 0$, alors d'après l'algorithme de couplage, on a forcément couplé P_1 avec une classe d'indice ℓ tel que $\ell \leq r_1$, et donc $\epsilon_{1,\ell} = 1$. Par ailleurs $m_1 = \min(0, r_1 - 1)$, et donc si $r_1 = 0$, $m_1 = -1$, et sinon $m_1 = 0$. Finalement on a bien

$$|P_1 \cap \hat{Q}_\ell| = t_1 + 1 + m_1.$$

- Hérité : on suppose que $\sum_{(i,j) \in \hat{\sigma}_{k-1}} |P_i \cap \hat{Q}_j| = \sum_{i=1}^{k-1} t_i + k - 1 + m_{k-1}$. Montrons que l'égalité est alors vérifiée à l'ordre k . L'hypothèse de récurrence implique, par le même raisonnement que précédemment, que :

$$k - 1 + m_{k-1} = \sum_{(i,j) \in \sigma_{k-1}} \epsilon_{i,j} = |\{(i,j) \in \sigma_{k-1}, j \leq r_i\}| \leq r_k.$$

On distingue alors à nouveau les deux cas suivants.

- Si $k - 1 - m_{k-1} = r_k$, toutes les classes \hat{Q}_j , $j \in \{1, \dots, r_k\}$ sont couplées avec une classe P_i , $i \leq k - 1$. L'algorithme couple alors P_k avec une classe \hat{Q}_ℓ telle que $\ell > r_k$, et donc $|P_k \cap \hat{Q}_\ell| = t_k$. Ainsi :

$$\sum_{(i,j) \in \hat{\sigma}_k} |P_i \cap \hat{Q}_j| = \sum_{(i,j) \in \hat{\sigma}_{k-1}} |P_i \cap \hat{Q}_j| + |P_k \cap \hat{Q}_\ell| = \sum_{i=1}^k t_i + k - 1 + m_{k-1}.$$

De plus $m_{k-1} = r_k - (k - 1)$, et donc

$$m_k = \min(m_{k-1}, r_k - k) = \min(m_{k-1}, m_{k-1} - 1) = m_{k-1} - 1.$$

D'où le résultat :

$$\sum_{(i,j) \in \hat{\sigma}_k} |P_i \cap \hat{Q}_j| = \sum_{i=1}^k t_i + k + m_k.$$

• Si $k - 1 + m_{k-1} < r_k$, alors l'algorithme couple P_k avec une classe \hat{Q}_ℓ telle que $\ell \leq r_k$ et donc $|P_k \cap \hat{Q}_\ell| = t_k + 1$. On a donc :

$$\sum_{(i,j) \in \hat{\sigma}_k} |P_i \cap \hat{Q}_j| = \sum_{(i,j) \in \hat{\sigma}_{k-1}} |P_i \cap \hat{Q}_j| + |P_k \cap \hat{Q}_\ell| = \sum_{i=1}^k t_i + k + m_{k-1}.$$

De plus $m_{k-1} \leq r_k - k$, et donc

$$m_k = \min(m_{k-1}, r_k - k) = m_{k-1}.$$

D'où à nouveau le résultat :

$$\sum_{(i,j) \in \hat{\sigma}_k} |P_i \cap \hat{Q}_j| = \sum_{i=1}^k t_i + k + m_k.$$

Ainsi cette proposition est vraie pour tout $k \in \{1, \dots, p\}$ et donc on obtient :

$$c_{\hat{\sigma}}(\mathcal{P}, \hat{\mathcal{Q}}) = \sum_{(i,j) \in \hat{\sigma}} |P_i \cap \hat{Q}_j| = \sum_{i=1}^p t_i + k + m_p = f(\mathcal{P}, q).$$

Finalement, comme

$$c(\mathcal{P}, \hat{\mathcal{Q}}) = \max_{\sigma} c_{\sigma}(\mathcal{P}, \hat{\mathcal{Q}}) \leq f(\mathcal{P}, q)$$

et que

$$c_{\hat{\sigma}}(\mathcal{P}, \hat{\mathcal{Q}}) = f(\mathcal{P}, q),$$

on en déduit la Proposition 1 :

$$c(\mathcal{P}, \hat{\mathcal{Q}}) = c_{\hat{\sigma}}(\mathcal{P}, \hat{\mathcal{Q}}) = f(\mathcal{P}, q).$$

□

LEMME 2. Pour toute partition \mathcal{P} à p classes et tout entier q tel que $p \leq q$, on a :

$$c_{\min}(\mathcal{P}, \leq q) \leq f(\mathcal{P}, q).$$

Proof. (Lemme 2). Ce lemme découle directement de la Proposition 1 :

$$c_{\min}(\mathcal{P}, \leq q) \leq c(\mathcal{P}, \hat{\mathcal{Q}}) = f(\mathcal{P}, q).$$

□

On aura aussi besoin des lemmes suivants :

LEMME 3. *On considère une partition \mathcal{P} sur X , un entier q et un élément $a \in X$. Alors :*

$$c_{\min}(\mathcal{P}, \leq q) \geq c_{\min}(\mathcal{P}_{\ominus} a, \leq q).$$

On rappelle que $\mathcal{P}_{\ominus} a$ dénote la partition sur $X \setminus \{a\}$ définie par :

- si $P_k = \{a\}$, $\mathcal{P}_{\ominus} a$ est constituée des classes P_i , $i \neq k$.
- sinon $\mathcal{P}_{\ominus} a$ est constituée des classes P_i , $i \neq k$ ainsi que de la classe $P_k \setminus \{a\}$.

Proof. (Lemme 3). Ce lemme se démontre de la même façon que le Lemme 1. On note \mathcal{P}' la partition $\mathcal{P}_{\ominus} a$. Soit \mathcal{Q} une partition sur X à au plus q classes et \mathcal{Q}' la partition $\mathcal{Q}_{\ominus} a$. Soit $\hat{\sigma}$ le couplage réalisant la concordance maximum entre \mathcal{P}' et \mathcal{Q}' . Par définition, puisque \mathcal{Q}' possède au plus q classes, on a

$$c_{\min}(\mathcal{P}', \leq q) \leq c(\mathcal{P}', \mathcal{Q}') = c_{\hat{\sigma}}(\mathcal{P}', \mathcal{Q}').$$

On note encore $\hat{\sigma}$ le couplage entre \mathcal{P} et \mathcal{Q} obtenu à partir de $\hat{\sigma}$ en remettant a à sa place dans chacune des partitions : si $P_k = \{a\}$, P_k est couplée avec une classe de \mathcal{Q} non couplée, ou s'il n'y en a pas, P_k n'est pas couplée. Alors par définition de la concordance :

$$c(\mathcal{P}, \mathcal{Q}) = \max_{\sigma} c_{\sigma}(\mathcal{P}, \mathcal{Q}) \geq c_{\hat{\sigma}}(\mathcal{P}, \mathcal{Q}).$$

Or on remarque l'inégalité

$$c_{\hat{\sigma}}(\mathcal{P}, \mathcal{Q}) \geq c_{\hat{\sigma}}(\mathcal{P}', \mathcal{Q}')$$

puisque soit $\hat{\sigma}$ associe les classes de \mathcal{P} et \mathcal{Q} contenant a et alors

$$c_{\hat{\sigma}}(\mathcal{P}, \mathcal{Q}) = c_{\hat{\sigma}}(\mathcal{P}', \mathcal{Q}') + 1,$$

soit ce n'est pas le cas et alors il y a égalité entre $c_{\hat{\sigma}}(\mathcal{P}, \mathcal{Q})$ et $c_{\hat{\sigma}}(\mathcal{P}', \mathcal{Q}')$.

Finalement on a :

$$c(\mathcal{P}, \mathcal{Q}) \geq c_{\hat{\sigma}}(\mathcal{P}, \mathcal{Q}) \geq c_{\hat{\sigma}}(\mathcal{P}', \mathcal{Q}') \geq c_{\min}(\mathcal{P}', \leq q),$$

et ce quelle que soit la partition \mathcal{Q} considérée. Donc pour toute partition \mathcal{Q} possédant au plus q classes, $c(\mathcal{P}, \mathcal{Q}) \geq c_{\min}(\mathcal{P}', \leq q)$, ce qui implique

$$c_{\min}(\mathcal{P}, \leq q) = \min_{\mathcal{Q} \text{ à } q \text{ classes}} c(\mathcal{P}, \mathcal{Q}) \geq c_{\min}(\mathcal{P}', \leq q).$$

□

LEMME 4. *Soient \mathcal{P} une partition sur X à p classes de cardinaux n_1, n_2, \dots, n_p et \mathcal{P}' une pseudo-partition sur $Y \subset X$ à p classes de cardinaux n'_1, n'_2, \dots, n'_p tels que $\forall i \in \{1, \dots, p\}$, $n_i \geq n'_i$. Soit $q \in \mathbb{N}$. Alors :*

$$c_{\min}(\mathcal{P}, \leq q) \geq c_{\min}(\mathcal{P}', \leq q).$$

Proof. (Lemme 4). Posons $k = \sum_{i=1}^p n_i - \sum_{i=1}^p n'_i$. On construit alors une suite $\mathcal{P}_0 = \mathcal{P}, \mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_k = \mathcal{P}'$ de pseudo-partitions définie de la façon suivante. Pour tout i compris entre 1 et k , soit a_i un élément d'une classe de \mathcal{P}_{i-1} tel que $a_i \in X \setminus Y$; on pose $\mathcal{P}_i = \mathcal{P}_{i-1} \ominus a_i$. D'après le lemme 3, on a donc $\forall i \in \{1, \dots, k\}$,

$$c_{\min}(\mathcal{P}_{i-1}, \leq q) \geq c_{\min}(\mathcal{P}_i, \leq q),$$

et donc

$$c_{\min}(\mathcal{P}, \leq q) \geq c_{\min}(\mathcal{P}', \leq q).$$

□

Proof. (Théorème 3). La preuve de ce théorème se trouve en Annexe. Elle est fondée sur la Proposition 1 et les Lemmes 2, 3 et 4 et sur un raisonnement par récurrence sur le nombre d'éléments n . □

4.2. CAS 2. VALEUR DE $c_{\min}(\mathcal{P}, \leq q)$ QUAND $p > q$

THÉORÈME 4. *Soit \mathcal{P} une partition à p classes sur X et q un entier vérifiant $q < p$. On note $n_i, i \in \{1, \dots, p\}$, les cardinaux des classes de \mathcal{P} et $r_i, i \in \{1, \dots, p\}$, les restes respectifs de la division entière des n_i par q . On suppose que les q plus grosses classes de \mathcal{P} sont indicées de 1 à q et dans l'ordre croissant des r_i . Alors*

$$c_{\min}(\mathcal{P}, \leq q) = \sum_{i=1}^q \left\lfloor \frac{n_i}{q} \right\rfloor + \min(0, \min_{1 \leq i \leq q} (r_i - i)) + q.$$

Remarque. Si on appelle \mathcal{P}' la partition constituée des q plus grosses classes de \mathcal{P} , le second terme de cette égalité correspond, d'après la notation introduite au cas 1, à $f(\mathcal{P}', q)$ et est égal, d'après le Théorème 3, à $c_{\min}(\mathcal{P}', \leq q)$.

Commençons par prouver la proposition suivante :

PROPOSITION 2. *Soient \mathcal{P} une partition à p classes définie sur X et un entier q tel que $p > q$. Soient \mathcal{P}' la partition constituée des q plus grosses classes de \mathcal{P} et $\hat{\mathcal{Q}}$ la pseudo-partition construite à partir de \mathcal{P} comme indiqué au début de la Section 4. On a :*

$$c(\mathcal{P}, \hat{\mathcal{Q}}) = f(\mathcal{P}', q).$$

Proof. (Proposition 2). On suppose que les classes de \mathcal{P} sont indicées comme décrit dans l'énoncé du Théorème 4. Soit un couplage $\hat{\sigma}$ optimal entre \mathcal{P} et $\hat{\mathcal{Q}}$ qui couple un maximum de classes de \mathcal{P} d'indices inférieurs à q .

• 1^{er} cas : $\hat{\sigma}$ couple les q premières classes de \mathcal{P} (et les q classes de $\hat{\mathcal{Q}}$). Alors les $p - q$ autres classes de \mathcal{P} ne servent à rien pour le couplage et donc :

$$c(\mathcal{P}, \hat{\mathcal{Q}}) = c_{\hat{\sigma}}(\mathcal{P}, \hat{\mathcal{Q}}) = c_{\hat{\sigma}}(\mathcal{P}', \hat{\mathcal{Q}}) = c(\mathcal{P}', \hat{\mathcal{Q}}).$$

En effet, s'il existait un couplage σ entre \mathcal{P}' et $\hat{\mathcal{Q}}$ meilleur que $\hat{\sigma}$, on pourrait trouver un couplage entre \mathcal{P} et $\hat{\mathcal{Q}}$ meilleur que $\hat{\sigma}$ en couplant les q premières classes de \mathcal{P}

avec celles de $\hat{\mathcal{Q}}$ selon σ et en ne couplant pas les $p - q$ autres classes de \mathcal{P} . Or d'après la Proposition 1, on a :

$$c(\mathcal{P}', \hat{\mathcal{Q}}) = f(\mathcal{P}', q),$$

et donc on obtient

$$c(\mathcal{P}, \hat{\mathcal{Q}}) = f(\mathcal{P}', q).$$

• 2^e cas : $\hat{\sigma}$ ne couple pas les q premières classes de \mathcal{P} . Alors, il existe $i \in \{1, \dots, q\}$ tel que P_i est non couplée et il existe j tel que \hat{Q}_j est non couplée avec les classes P_1, \dots, P_q . La classe \hat{Q}_j est donc couplée avec P_k , $k > q$, ou n'est pas couplée. On découple alors \hat{Q}_j de P_k si nécessaire et on la couple avec P_i . Par construction de $\hat{\mathcal{Q}}$, $|P_k \cap \hat{Q}_j| \leq |P_i \cap \hat{Q}_j|$ et donc on n'y perd pas en concordance mais la nouvelle affectation (donc optimale) couple davantage de classes de \mathcal{P} d'indices compris entre 1 et q , ce qui entraîne une contradiction. Le second cas est donc impossible. \square

Proof. (Théorème 4). D'après le Lemme 4, on a :

$$c_{\min}(\mathcal{P}', \leq q) \leq c_{\min}(\mathcal{P}, \leq q).$$

D'autre part, on a par définition :

$$c_{\min}(\mathcal{P}, \leq q) \leq c(\mathcal{P}, \hat{\mathcal{Q}}).$$

Or, d'après la Proposition 2 et le Théorème 3, $c(\mathcal{P}, \hat{\mathcal{Q}}) = f(\mathcal{P}', q) = c_{\min}(\mathcal{P}', \leq q)$, et donc

$$c_{\min}(\mathcal{P}, \leq q) \leq c_{\min}(\mathcal{P}', \leq q).$$

D'où finalement l'égalité énoncée dans le Théorème 4. \square

Nous avons ainsi déterminé, pour tout entier q , la valeur de $c_{\min}(\mathcal{P}, \leq q)$ (cf. Théorèmes 3 et 4).

Remarque. Dans les Théorèmes 3 et 4, seuls le nombre de classes et les cardinaux des classes de \mathcal{P} entrent en compte : la répartition des éléments dans les classes de \mathcal{P} n'influe pas sur la valeur de $c_{\min}(\mathcal{P}, \leq q)$.

On peut de plus déduire des Propositions 1 et 2 le Corollaire suivant.

COROLLAIRE 1. *Pour toute partition \mathcal{P} et tout entier positif q , la pseudo-partition $\hat{\mathcal{Q}}$ au début de la Section 4 est, parmi les partitions à au plus q classes, une des moins concordantes avec \mathcal{P} :*

$$c_{\min}(\mathcal{P}, \leq q) = c(\mathcal{P}, \hat{\mathcal{Q}}).$$

Enfin, la proposition suivante établit un lien entre $c_{\min}(\mathcal{P}, \leq q)$, dont on connaît maintenant la valeur, et $c_{\min}(\mathcal{P}, q)$.

PROPOSITION 3. *Soit \mathcal{P} une partition sur X à p classes de cardinaux n_1, n_2, \dots, n_p et q un entier positif. On a alors :*

$$c_{\min}(\mathcal{P}, \leq q) = c_{\min}(\mathcal{P}, \min(q, \max_{1 \leq i \leq p} n_i)).$$

Proof. (Proposition 3). On pose $q' = \min(q, \max_{1 \leq i \leq p} n_i)$.

Tout d'abord, puisque $q' \leq q$, on a

$$c_{\min}(\mathcal{P}, \leq q) \leq c_{\min}(\mathcal{P}, q').$$

D'après le Corollaire 1, on a : $c_{\min}(\mathcal{P}, \leq q) = c(\mathcal{P}, \hat{\mathcal{Q}})$.

Soit $\hat{\mathcal{Q}}'$ la partition composée des q' premières classes de $\hat{\mathcal{Q}}$. Si $q' = q$, $\hat{\mathcal{Q}}' = \hat{\mathcal{Q}}$. Sinon, d'après la construction de $\hat{\mathcal{Q}}$, seules ses q' premières classes sont non vides. Ainsi dans tous les cas on a $c(\mathcal{P}, \hat{\mathcal{Q}}) = c(\mathcal{P}, \hat{\mathcal{Q}}')$ et donc :

$$c_{\min}(\mathcal{P}, \leq q) = c(\mathcal{P}, \hat{\mathcal{Q}}').$$

Enfin, puisque $\hat{\mathcal{Q}}'$ est une partition possédant exactement q' classes, on a :

$$c_{\min}(\mathcal{P}, q') \leq c(\mathcal{P}, \hat{\mathcal{Q}}').$$

Finalement on a :

$$c(\mathcal{P}, \hat{\mathcal{Q}}') = c_{\min}(\mathcal{P}, \leq q) \leq c_{\min}(\mathcal{P}, q') \leq c(\mathcal{P}, \hat{\mathcal{Q}}'),$$

et donc

$$c_{\min}(\mathcal{P}, \leq q) = c_{\min}(\mathcal{P}, q') = c_{\min}(\mathcal{P}, \min(q, \max_{1 \leq i \leq p} n_i)).$$

□

On peut ainsi calculer $c_{\min}(\mathcal{P}, q)$ si $q \leq \max_{1 \leq i \leq p} n_i$. Si $q > \max_{1 \leq i \leq p} n_i$, la construction optimale proposée ne peut plus être appliquée ; une construction optimale pour ce cas semble très dépendante des cardinaux des classes de la partition initiale et est donc sans doute difficile à définir pour le cas général. On peut toutefois remarquer que si $c_{\min}(\mathcal{P}, \leq q)$ est différent de $c_{\min}(\mathcal{P}, \leq q - 1)$, alors il y a égalité entre $c_{\min}(\mathcal{P}, q)$ et $c_{\min}(\mathcal{P}, \leq q)$.

5. CONCLUSION

En guise de conclusion à cet article, on résume les résultats trouvés en termes de distance de transfert. Soient X un ensemble fini à n éléments et \mathcal{P} une partition sur X à p classes.

La distance maximum $\theta_{\max}(\mathcal{P})$ entre \mathcal{P} et une partition quelconque sur X est donnée par :

$$\theta_{\max}(\mathcal{P}) = n - p - \min[0, \min_{1 \leq i \leq p} (n_i - i)],$$

où n_1, n_2, \dots, n_p désignent les cardinaux des classes de \mathcal{P} ordonnés dans l'ordre croissant (cf. Théorème 2).

La distance maximum $\theta_{\max}(\mathcal{P}, \leq q)$ entre \mathcal{P} et une partition à au plus q classes est donnée par :

- si $p \leq q$,

$$\theta_{\max}(\mathcal{P}, \leq q) = n - \sum_{i=1}^p \left\lfloor \frac{n_i}{q} \right\rfloor - \min(0, \min_{1 \leq i \leq p} (r_i - i)) - p$$

où n_1, n_2, \dots, n_p désignent les cardinaux des classes de \mathcal{P} , r_1, r_2, \dots, r_p désignent les restes de la division entière des n_i par q et on suppose que $r_1 \leq r_2 \leq \dots \leq r_p$ (cf. Théorème 3).

- si $p > q$,

$$\theta_{\max}(\mathcal{P}, \leq q) = n - \sum_{i=1}^q \left\lfloor \frac{n_i}{q} \right\rfloor - \min(0, \min_{1 \leq i \leq q} (r_i - i)) - q$$

où n_1, n_2, \dots, n_p désignent les cardinaux des classes de \mathcal{P} , r_1, r_2, \dots, r_p désignent les restes de la division entière des n_i par q et on suppose que les q plus grosses classes de \mathcal{P} sont indicées de 1 à q et dans l'ordre croissant des r_i (cf. Théorème 4).

La distance maximum $\theta_{\max}(\mathcal{P}, q)$ entre \mathcal{P} et une partition à exactement q classes est donnée, si $q \leq \max_{1 \leq i \leq p} n_i$, par :

$$\theta_{\max}(\mathcal{P}, q) = \theta_{\max}(\mathcal{P}, \leq q)$$

où n_1, n_2, \dots, n_p désignent les cardinaux des classes de \mathcal{P} .

Remerciements. Ce travail a bénéficié du soutien du CNRS, dans le cadre de l'ACI IMP-Bio, grâce à Alain Guénoche (CNRS) que nous remercions. Nous remercions aussi les rapporteurs anonymes pour leurs remarques constructives qui ont permis d'améliorer la présentation du texte.

BIBLIOGRAPHIE

AHUJA R.K., MAGNANTI T.L., ORLIN J.B., *Network flows*, Englewood Cliffs (NJ), Prentice Hall, 1993.

ALPERT C.J., KANG, A., "Recent direction in netlist partitioning: a survey", *Integration: the VLSI Journal* 19, 1-2, 1995, p. 1-81.

ARABIE P., HUBERT L.J., DE SOETE G., *Clustering and classification*, River Edge (NJ), Eds. World Scientific Publishers, 1996.

BATAGELJ V., MRVRAR M., ZVERSNIK M., "Partitioning approach to visualisation of large graphs", *Lecture Notes in Computer Science* 1731, Springer, 1999, p. 90-97.

CHARON I., DENEUD L., GUÉNOCHE A., HUDRY O., "Maximum transfer distance between partitions", *Journal of Classification* 23, 1, 2006, p. 103-121.

DAY W., "The complexity of computing metric distances between partitions", *Mathematical Social Sciences* 1, 1981, p. 269-287.

- FRAYSSEIX (de) H., KUNTZ P., "Pagination of large scale networks; embedding a graphe in \mathbb{R}^n for effective partitioning", *Algorithmic review* 2(3), 1992, p. 105-112.
- DENËUD L., GUÉNOCHE A., "Comparison of distance indices between partitions", V. Batagelj et al. (eds.), *Proceedings of IFCS'2006, Data Science and Classification*, Ljubljana (Slovenia), Springer, July 2006, p. 21-28.
- GETZ G., LEVINE E., DOMANY E., "Coupled two-way clustering analysis of gene microarray data", *Proc. Natl. Acad. Sci. USA* 97 22, 2000, p. 12079-12084.
- GUÉNOCHE A., "Clustering by graph density", *Proceedings of the International Federation of the Classification Societies*, D. Banks and al. (eds), Springer, 2004, p. 15-23.
- GUÉNOCHE A., "Comparison of algorithms in graph partitioning", ALIO/EURO Conference on Combinatorial Optimization, Paris, October 2005.
- HUBERT L., ARABIE P., "Comparing partitions", *J. of Classification* 2, 1985, p. 193-218.
- KUHN H.W., "The Hungarian method for the assignment problem", *Naval Res. Logist. Quart.* 2, 1955, p. 83-97.
- KUHN H.W., "Variants on the Hungarian method for the assignment problems", *Naval Res. Logist. Quart.* 3, 1956, p. 253-258.
- LERMAN I.C., "Comparing partitions (mathematical and statistical aspects)", *Classification and Related Methods of Data Analysis*, H.H Bock (Ed.) Elsevier Science Publishers, 1988, p. 121-131.
- LOVASZ L., PLUMMER M., "Matching Theory", *Annals of Discrete Mathematics* 29, Amsterdam, North Holland, 1986.
- MATSUDA H., ISHIHARA T., HASHIMOTO A., "Classifying molecular sequences using a linkage graph with their pairwise similarities", *Theoretical Computer Science* 210, 1999, p. 305-325.
- MOODY J., "Identifying dense clusters in large networks", *Social Networks* 23, 2001, p. 261-283.
- NEWMAN E.J., "The structure of scientific collaboration networks", *Proc. Natl. Acad. Sci. USA* 98, 2001, p. 404-409.
- RAND W.M., "Objective criteria for the evaluation of clustering methods", *J. Amer. Statist. Assoc.* 66, 1971, p. 846-850.
- RÉGNIER S., "Quelques aspects mathématiques des problèmes de classification automatique", I.C.C. Bulletin 4, 1965. [Réédité dans *Mathématiques et Sciences humaines* 82, 1983, p. 13-29].
- SEIDMAN S.B., "Network structure and minimum degree", *Social Networks* 5, 1983, p. 269-287.
- van DONGEN S., *Graph clustering by flow simulation*, thèse de doctorat, University of Utrecht, 2000.

ANNEXE

PREUVE DU THÉORÈME 3. On raisonne par récurrence sur le nombre d'éléments n .

- Initialisation : si $n = 1$, il est clair que la formule est vraie. En effet on a $c_{\min}(\mathcal{P}, \leq q) = 1 = f(\mathcal{P}, q)$.

- Hérité : on suppose maintenant que le théorème est vrai pour toute partition sur un ensemble possédant strictement moins de n éléments, $n \geq 2$. On considère une partition \mathcal{P} à p classes sur n éléments. On utilise les mêmes notations que précédemment : pour tout $i \in \{1, \dots, p\}$, n_i désigne le cardinal de la classe P_i de \mathcal{P} , et on définit les entiers naturels t_i et r_i par $n_i = t_i q + r_i$, avec $r_i \in \{0, \dots, q-1\}$. On distingue alors différents cas :

- **cas 1** : $\exists i_0 \in \{1, \dots, p\}$ tel que $r_{i_0} - i_0 \geq 1$

On considère un élément a appartenant à P_{i_0} et la partition $\mathcal{P}' = \mathcal{P} \ominus a$.

EXEMPLE 4. Pour illustrer ce cas, considérons la partition $\mathcal{P} = (123|456789)$ sur l'ensemble à 9 éléments et $q=4$. On a $n_1 = 3$ et $n_2 = 6$, ce qui donne $r_1 = 3$ et $r_2 = 2$. On réindice les classes par ordre croissant des r_i ; on obtient $r_i - i = 1$, $i \in \{1, 2\}$. On choisit alors arbitrairement $i_0 = 1$, puisque ici les deux indices pourraient convenir ; cet indice correspond à la classe à 6 éléments.

Construisons maintenant la partition $\hat{\mathcal{Q}}$ associée à \mathcal{P} : $\hat{\mathcal{Q}} = (148|259|36|7)$. La valeur de la concordance entre ces deux partitions, qui peut se calculer par le calcul de $f(\mathcal{P}, 4)$, vaut 3, et correspond à l'appariement de P_2 avec une des deux premières classes de $\hat{\mathcal{Q}}$, et P_1 avec une des classes restantes parmi les trois premières.

On considère maintenant une nouvelle partition \mathcal{P}' construite à partir de \mathcal{P} en supprimant un des éléments de la classe i_0 : $\mathcal{P}' = (123|45678)$ (on choisit de supprimer par exemple l'élément 9). La partition $\hat{\mathcal{Q}}'$ est alors construite comme suit : $\hat{\mathcal{Q}}' = (148|25|36|7)$. Un couplage optimal entre les classes de \mathcal{P}' et celles de $\hat{\mathcal{Q}}'$ doit associer P'_2 avec $\hat{\mathcal{Q}}'_1$ et P'_1 avec une des deux classes suivantes. Ce couplage donne aussi une concordance égale à 3.

On remarque sur cet exemple que, dans ce cas, la suppression d'un élément de la classe d'indice i_0 ne modifie pas la valeur de $f(\mathcal{P}, q)$. C'est sur cette observation qu'est basée la preuve qui suit.

La partition \mathcal{P}' possède p classes dont toutes excepté P'_{i_0} sont identiques aux classes correspondantes de \mathcal{P} . On note n'_i le cardinal de la classe i de \mathcal{P}' et on pose $n'_i = t'_i q + r'_i$, $r'_i \in \{0, \dots, q-1\}$. On a alors $\forall i \in \{1, \dots, p\}$, $i \neq i_0$, $r'_i = r_i$ et $t'_i = t_i$, et $r'_{i_0} = r_{i_0} - 1$. Or puisque $r_{i_0} > 1$, on a donc aussi $t'_{i_0} = t_{i_0}$. Donc :

$$f(\mathcal{P}', q) = \sum_{i=1}^p t'_i + \min(0, \min_{1 \leq i \leq p} (r'_i - i)) + p = \sum_{i=1}^p t_i + \min(0, \min_{1 \leq i \leq p, i \neq i_0} (r_i - i), r'_{i_0} - i_0) + p.$$

Puisque $r'_{i_0} - i_0 \geq 0$,

$$f(\mathcal{P}', q) = \sum_{i=1}^p t_i + \min(0, \min_{1 \leq i \leq p, i \neq i_0} (r_i - i)) + p.$$

De plus, puisque $r_{i_0} - i_0 \geq 1$,

$$f(\mathcal{P}', q) = \sum_{i=1}^p t_i + \min(0, \min_{1 \leq i \leq p} (r_i - i)) + p = f(\mathcal{P}, q).$$

On applique maintenant l'hypothèse de récurrence puisque \mathcal{P}' est une partition sur $n - 1$ éléments :

$$c_{\min}(\mathcal{P}', \leq q) = f(\mathcal{P}', q) = f(\mathcal{P}, q).$$

De plus d'après le Lemme 3

$$c_{\min}(\mathcal{P}', \leq q) \leq c_{\min}(\mathcal{P}, \leq q), \text{ et donc } f(\mathcal{P}, q) \leq c_{\min}(\mathcal{P}, \leq q).$$

On sait par ailleurs par le Lemme 2 que

$$c_{\min}(\mathcal{P}, \leq q) \leq f(\mathcal{P}, q),$$

d'où le résultat :

$$f(\mathcal{P}, q) = c_{\min}(\mathcal{P}, \leq q).$$

- cas 2 : $\forall i \in \{1, \dots, p\}, r_i - i \leq 0$

Soit i_0 tel que $r_{i_0} - i_0 = \min_{1 \leq i \leq p} (r_i - i)$. On distingue alors à nouveau deux cas :

- cas 2.1 : $\exists j \in \{1, \dots, p\}$ tel que $r_j \neq 0$ et $r_j - j > r_{i_0} - i_0$

On considère alors un élément $a \in P_j$ et la partition $\mathcal{P}' = \mathcal{P} \ominus a$.

EXEMPLE 5. Soit $\mathcal{P} = (1, 2, 3, 4|5, 6)$ et $q=4$. On a $n_1 = 4$ et $n_2 = 2$, soit $r_1 = 0$ et $r_2 = 2$, ce qui donne $i_0=1$ (on remarque que $r_i - i \leq 0, i \in \{1, 2\}$) et $j=2$ ($r_2 \neq 0$ et $r_2 - 2 > r_1 - 1$). La partition $\hat{\mathcal{Q}}$ vaut $(1, 5|2, 6|3|4)$. La concordance entre \mathcal{P} et $\hat{\mathcal{Q}}$ vaut donc 2 et correspond au couplage de P_2 avec une des deux premières classes de $\hat{\mathcal{Q}}$ et de P_1 avec n'importe quelle classe restante.

Si maintenant on considère la partition $\mathcal{P}' = (1, 2, 3, 4|5)$ (on a supprimé l'élément 6 de la classe $j = 2$), on obtient $\hat{\mathcal{Q}}' = (1, 5|2|3|4)$. Un meilleur couplage entre les classes de \mathcal{P}' et celles de $\hat{\mathcal{Q}}'$ associe forcément P'_2 avec $\hat{\mathcal{Q}}'_1$ et P'_1 avec n'importe quelle classe restante. On obtient donc aussi $c(\mathcal{P}', \hat{\mathcal{Q}}') = 2$.

On remarque donc à nouveau que dans ce cas, la suppression d'un élément de \mathcal{P} ne modifie pas la valeur de $f(\mathcal{P}, q)$. C'est par ce raisonnement qu'on traite le cas 2.1.

La partition \mathcal{P}' possède p classes dont toutes excepté P'_j sont identiques aux classes correspondantes de \mathcal{P} . De plus $t'_j = t_j$ puisque $r_j \neq 0$, et donc :

$$\begin{aligned} f(\mathcal{P}', q) &= \sum_{i=1}^p t'_i + \min(0, \min_{1 \leq i \leq p} (r'_i - i)) + p = \sum_{i=1}^p t_i + \min(0, \min_{1 \leq i \leq p, i \neq j} (r_i - i), r'_j - j) + p \\ &= \sum_{i=1}^p t_i + \min(0, r_{i_0} - i_0, r'_j - j) + p \end{aligned}$$

car $i_0 \neq j$. Or $r'_j - j = r_j - 1 - j \geq r_{i_0} - i_0$, et donc :

$$f(\mathcal{P}', q) = \sum_{i=1}^p t_i + r_{i_0} - i_0 + p = f(\mathcal{P}, q).$$

Finalement en appliquant l'hypothèse de récurrence puisque \mathcal{P}' est une partition sur $n - 1$ éléments, et les lemmes 2 et 3, on obtient de la même façon que dans le cas 1 :

$$f(\mathcal{P}, q) = c_{\min}(\mathcal{P}, \leq q).$$

- **cas 2.2** : $\forall j \in \{1, \dots, p\}$, $r_j = 0$ ou $r_j - j = r_{i_0} - i_0$

Ce cas correspond à la configuration :

$$r_1 = r_2 = \dots = r_k = 0, r_{k+1} = 1, r_{k+2} = 2, \dots, r_p = p - k$$

pour un certain k compris entre 0 et p .

EXEMPLE 6. La partition $\mathcal{P} = (1, 2, 3|4, 5, 6, 7|8, 9)$ avec $q = 3$ entre dans ce cas. On a en effet $r_1 = 0$, $r_2 = 1$ et $r_3 = 2$. Construisons la partition $\hat{\mathcal{Q}}$ correspondante : $\hat{\mathcal{Q}} = (1, 4, 7, 8|2, 5, 9|3, 6)$. Ici le meilleur couplage entre ces deux partitions est unique et associe forcément P_2 avec \hat{Q}_1 , P_3 avec \hat{Q}_2 , et P_1 avec \hat{Q}_3 . Il donne une concordance égale à 4.

On voit sur cet exemple que la technique utilisée plus haut consistant à retirer un élément n'est plus applicable, car elle produirait une diminution de la valeur de la concordance. On doit donc procéder différemment, en considérant à nouveau divers cas.

Nous commencerons par traiter le cas $k \geq 1$ puis nous montrerons que le cas $k = 0$ s'y ramène.

- **cas 2.2.1** : $k \geq 1$

Soit \mathcal{Q} une pseudo-partition quelconque à q classes. On veut montrer l'inégalité $c(\mathcal{P}, \mathcal{Q}) \geq f(\mathcal{P}, q)$.

Tout d'abord remarquons que dans ce cas $\min(0, \min_{1 \leq i \leq p}(r_i - i)) = -k$, et donc

$$f(\mathcal{P}, q) = \sum_{i=1}^p t_i + p - k.$$

On définit ϵ_i , $i \in \{1, \dots, p\}$, par : $\epsilon_i = 0$ si $1 \leq i \leq k$ et $\epsilon_i = 1$ si $k < i \leq p$. On définit alors le couple d'indices (i_0, j_0) de la façon suivante :

- $|P_{i_0} \cap Q_{j_0}| - (t_{i_0} + \epsilon_{i_0}) = \max_{(i,j)}(|P_i \cap Q_j| - (t_i + \epsilon_i))$
- de préférence $i_0 > k$ et, dans ce cas, i_0 est le plus petit possible.

On construit alors deux pseudo-partitions \mathcal{P}' et \mathcal{Q}' à partir de \mathcal{P} et \mathcal{Q} en retirant les éléments de l'ensemble $P_{i_0} \cup Q_{j_0}$. Ces deux pseudo-partitions possèdent respectivement $p-1$ et $q-1$ classes (éventuellement vides). On note $P'_1, \dots, P'_{i_0-1}, P'_{i_0+1}, \dots, P'_p$ les classes de \mathcal{P}' . On a alors :

$$\forall i \in \{1, \dots, p\}, i \neq i_0, n'_i = n_i - |P_i \cap Q_{j_0}|$$

et donc

$$c(\mathcal{P}, \mathcal{Q}) \geq c_{\min}(\mathcal{P}, \leq q) \geq c_{\min}(\mathcal{P}', \leq q-1) + |P_{i_0} \cap Q_{j_0}|. \quad (\star)$$

On pose

$$\alpha = |P_{i_0} \cap Q_{j_0}| - (t_{i_0} + \epsilon_{i_0})$$

et on considère deux entiers s et r avec $r \in \{0, \dots, q-2\}$ tels que $\alpha = (q-1)s + r$. D'après la définition des indices i_0 et j_0 , on sait que $\forall (i, j) \in \{1, \dots, p\} \times \{1, \dots, q\}$, $|P_i \cap Q_j| - (t_i + \epsilon_i) \leq |P_{i_0} \cap Q_{j_0}| - (t_{i_0} + \epsilon_{i_0})$. En particulier pour $j = j_0$, on a $\forall i \in \{1, \dots, p\}$, $|P_i \cap Q_{j_0}| - (t_i + \epsilon_i) \leq |P_{i_0} \cap Q_{j_0}| - (t_{i_0} + \epsilon_{i_0})$, ce qui équivaut à :

$$\forall i \in \{1, \dots, p\}, \quad |P_i \cap Q_{j_0}| \leq \alpha + t_i + \epsilon_i.$$

Nous distinguons alors à nouveau deux cas :

- **cas 2.2.1.1** : $i_0 \leq k$

Tout d'abord, notons que dans ce cas $\epsilon_{i_0} = 0$ et donc

$$|P_{i_0} \cap Q_{j_0}| = \alpha + t_{i_0}.$$

Puisque i_0 est choisi si possible tel que $i_0 > k$, alors $i_0 \leq k$ implique que $\forall i > k$, $|P_i \cap Q_{j_0}| < \alpha + t_i + \epsilon_i$. Dans ce cas on a $\epsilon_i = 1$ et donc on obtient $|P_i \cap Q_{j_0}| \leq \alpha + t_i$. Si $i \leq k$, on a $\epsilon_i = 0$ et donc on obtient aussi $|P_i \cap Q_{j_0}| \leq \alpha + t_i$. Finalement

$$\forall i \in \{1, \dots, p\}, \quad |P_i \cap Q_{j_0}| \leq \alpha + t_i.$$

En remplaçant maintenant $|P_i \cap Q_{j_0}|$ par $n_i - n'_i$ et α par $(q-1)s + r$, on obtient :

$$\forall i \in \{1, \dots, p\}, \quad n'_i \geq n_i - (q-1)s - r - t_i;$$

on remplace alors n_i par $qt_i + r_i$:

$$\forall i \in \{1, \dots, p\}, \quad n'_i \geq (q-1)(t_i - s) + r_i - r.$$

On définit alors les entiers n''_i , $i \in \{1, \dots, p\}$, par :

$$n''_i = (q-1)(t_i - s) + r_i - r.$$

Soient $I^+ = \{i \in \{1, \dots, p\}, i \neq i_0, n''_i \geq 0\}$ et $I^- = \{i \in \{1, \dots, p\}, i \neq i_0, n''_i < 0\}$. On note de plus $h_1 = |\{i \in I^-, r_i < r\}|$ et $h_2 = |\{i \in I^-, r_i \geq r\}|$.

On considère maintenant la pseudo-partition \mathcal{P}'' à $|I^+|$ classes de cardinaux n''_i , $i \in I^+$. Pour $i \in I^+$, on pose $n''_i = t''_i(q-1) + r''_i$, avec $r''_i \in \{0, \dots, q-2\}$. On remarque que :

• si $r_i - r < 0$,

$$\begin{cases} t''_i &= t_i - s - 1 \\ r''_i &= r_i - r + q - 1 \end{cases}$$

• si $0 \leq r_i - r \leq q-2$,

$$\begin{cases} t''_i &= t_i - s \\ r''_i &= r_i - r \end{cases}$$

• si $r_i - r = q-1$,

$$\begin{cases} t''_i &= t_i - s + 1 \\ r''_i &= 0. \end{cases}$$

Par définition, $\forall i \in \{1, \dots, p\}$, $n'_i \geq n''_i$, ce qui implique d'après le Lemme 4 que

$$c_{\min}(\mathcal{P}', \leq q-1) \geq c_{\min}(\mathcal{P}'', \leq q-1). \quad (\diamond)$$

Calculons un minorant de $c_{\min}(\mathcal{P}'', \leq q-1)$ en appliquant l'hypothèse de récurrence. On doit donc calculer t''_i et $\min(0, \min_{i \in I^+}(u_i - i))$, où (u_i) est la suite (r''_i) ordonnée dans l'ordre croissant. Pour cela étudions la suite des $r_i - r$ pour $i \in I^+$.

Pour $i \in \{1, \dots, p\}$, $i \neq i_0$, la suite des $r_i - r$ s'écrit :

$$\underbrace{-r, -r, \dots, -r}_{k-1 \text{ fois}}, 1-r, 2-r, \dots, p-k-r.$$

La suite restreinte à I^+ est donc de la forme :

$$\underbrace{-r, -r, \dots, -r}_{\leq k-1 \text{ termes}}, \underbrace{\beta_1 - r, \beta_2 - r, \dots, \beta_l - r}_{l \leq p-k \text{ termes}}$$

pour des β_i appropriés avec $\beta_i > 0$, $\beta_{i+1} > \beta_i$ et $\beta_l \leq p - k$. Considérons maintenant les trois cas suivants :

2.2.1.1.1 $r \neq 0$ et $p - k - r \geq 0$

• si $\beta_l - r \geq 0$, soit il existe un indice j tel que $\beta_j - r < 0$ et $\beta_{j+1} - r \geq 0$, soit on a $\beta_i - r \geq 0$ pour tout $i \in I^+$. Dans la suite $(r_i - r)_{i \in \{1, \dots, p\}, i \neq i_0}$, il y a $p - k - r + 1$ termes positifs ou nuls et $k + r - 2$ termes strictement négatifs. Dans la suite réduite à I^+ , il y a donc $p - k - r + 1 - h_2$ termes positifs ou nuls et $k + r - 2 - h_1$ termes strictement négatifs. La suite des r_i'' ordonnée, notée $u_1, u_2, \dots, u_{|I^+|}$, s'écrit donc de la façon suivante :

$$\leq k - 1 \text{ termes}$$

$$\underbrace{\beta_{j+1} - r, \beta_{j+2} - r, \dots, \beta_l - r}_{p - k - r + 1 - h_2 \text{ termes}}, \underbrace{-r + q - 1, \dots, -r + q - 1, \beta_1 - r + q - 1, \dots, \beta_j - r + q - 1}_{k + r - 2 - h_1 \text{ termes}}.$$

En effet $\beta_l - r \leq -r + q - 1$ car $p - k - r \leq -r + q - 1$ ($k \geq 1$). Le minimum des $(u_i - i)$ est donc atteint soit pour $i = 1$ soit pour le plus grand i , noté \hat{i} , tel que $u_i = -r + q - 1$. En effet la suite des $(u_i - i)$ est tout d'abord croissante puisque $\beta_{i+1} > \beta_i$, elle est ensuite décroissante à partir du premier terme de (u_i) valant $-r + q - 1$ et jusqu'à \hat{i} ; enfin la dernière partie de la suite est à nouveau croissante à partir de \hat{i} .

Or puisque $u_1 = \beta_{j+1} - r \geq 0$, alors $u_1 - 1 \geq -1$.

On sait de plus que $\hat{i} \leq p - k - r + 1 - h_2 + (k - 1) \leq p - r$, ce qui donne $u_{\hat{i}} - \hat{i} \geq -r + q - 1 - (p - r) = q - p - 1 \geq -1$. Ainsi

$$\min_{i \in I^+} (u_i - i) = \min(u_1 - 1, u_{\hat{i}} - \hat{i}) \geq -1.$$

• si $\beta_l - r < 0$, tous les termes de la suite $(r_i - r)$ pour $i \in I^+$ sont strictement négatifs et donc $\forall i \in I^+, r_i'' = r_i - r + q - 1$. La suite des r_i'' est donc ordonnée comme celle des r_i et s'écrit :

$$\underbrace{-r + q - 1, \dots, -r + q - 1}_{\leq k-1 \text{ fois}}, \beta_1 - r + q - 1, \beta_2 - r + q - 1, \dots, \beta_l - r + q - 1.$$

Le minimum des $r_i'' - i$, pour $i \in I^+$, est donc atteint pour le plus grand i tel que $r_i'' = -r + q - 1$, c'est-à-dire pour $i \leq k - 1$, donc

$$\min_{i \in I^+} (r_i'' - i) \geq -r + q - k.$$

Or nous avons supposé dans ce cas que $p - k - r \geq 0$, ce qui implique que $q - k - r \geq 0$, et donc nous avons aussi :

$$\min_{i \in I^+} (u_i - i) = \min_{i \in I^+} (r_i'' - i) \geq -1.$$

On peut maintenant calculer un minorant de $c_{\min}(\mathcal{P}'', \leq q-1)$. D'après l'hypothèse de récurrence :

$$c_{\min}(\mathcal{P}'', \leq q-1) = \sum_{i \in I^+} t_i'' + \min(0, \min_{i \in I^+}(u_i - i)) + |I^+|.$$

En appliquant les résultats précédents, on obtient alors :

$$c_{\min}(\mathcal{P}'', \leq q-1) \geq \sum_{i \in I^+, r_i - r \geq 0} (t_i - s) + \sum_{i \in I^+, r_i - r < 0} (t_i - s - 1) - 1 + |I^+|.$$

De plus puisqu'il y a $k + r - 2 - h_1$ termes dans la seconde somme de cette expression,

$$c_{\min}(\mathcal{P}'', \leq q-1) \geq \sum_{i \in I^+} (t_i - s) - k - r + 2 + h_1 - 1 + |I^+|.$$

Revenons maintenant aux partitions \mathcal{P} et \mathcal{Q} . D'après (\star) et (\diamond) , on a

$$c(\mathcal{P}, \mathcal{Q}) \geq c_{\min}(\mathcal{P}'', \leq q-1) + |P_{i_0} \cap Q_{j_0}|.$$

Développons cette inégalité. D'après la minoration de $c_{\min}(\mathcal{P}'', \leq q-1)$ établie précédemment, il vient

$$c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i \in I^+} (t_i - s) - k - r + 2 + h_1 - 1 + |I^+| + |P_{i_0} \cap Q_{j_0}|.$$

On rappelle (cf. début du cas 2.2.1.1) que $|P_{i_0} \cap Q_{j_0}| = \alpha + t_{i_0} = (q-1)s + r + t_{i_0}$. Ainsi :

$$\begin{aligned} c(\mathcal{P}, \mathcal{Q}) &\geq \sum_{i \in I^+} (t_i - s) - k - r + 2 + h_1 - 1 + |I^+| + (q-1)s + r + t_{i_0} \\ &\Leftrightarrow c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i \in I^+} t_i - \sum_{i \in I^+} s + (q-1)s - k + 1 + h_1 + |I^+| + t_{i_0} \\ &\Leftrightarrow c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i \in I^+ \cup \{i_0\}} t_i - |I^+|s + (q-p)s + (p-1)s - k + 1 + h_1 + |I^+|, \end{aligned}$$

et, puisque $q \geq p$,

$$c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i \in I^+ \cup \{i_0\}} t_i - |I^+|s + (p-1)s - k + 1 + h_1 + |I^+|.$$

Avec $|I^+| + |I^-| = p - 1$, il vient :

$$\begin{aligned} c(\mathcal{P}, \mathcal{Q}) &\geq \sum_{i \in I^+ \cup \{i_0\}} t_i - |I^+|s + (|I^+| + |I^-|)s - k + 1 + h_1 + (p-1 - |I^-|) \\ &\Leftrightarrow c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i \in I^+ \cup \{i_0\}} t_i + |I^-|s - k + h_1 - |I^-| + p. \end{aligned}$$

En utilisant maintenant $|I^-| = h_1 + h_2$, on obtient :

$$c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i \in I^+ \cup \{i_0\}} t_i + |I^-|s - k - h_2 + p.$$

On cherche maintenant à minorer s . Revenons à la définition de n_i'' :

$$n_i'' = (q-1)(t_i - s) + r_i - r.$$

Or puisque $r_i \in \{0, \dots, q-1\}$ et $r \in \{0, \dots, q-2\}$, on a $r_i - r \geq 2 - q$ et donc

$$n_i'' \geq (q-1)(t_i - s) - q + 2 \geq (q-1)(t_i - s - 1) + 1.$$

Donc si $n_i'' < 0$, on a

$$t_i - s - 1 \leq -1 \Leftrightarrow s \geq t_i.$$

Si de plus $r_i - r \geq 0$, $n_i'' = (q-1)(t_i - s) + r_i - r < 0$ implique que $t_i - s < 0$, d'où $s \geq t_i + 1$. Ainsi

$$\begin{aligned} |I^-|s &= \sum_{i \in I^-} s = \sum_{i \in I^-, r_i < r} s + \sum_{i \in I^-, r_i \geq r} s \geq \sum_{i \in I^-, r_i < r} t_i + \sum_{i \in I^-, r_i \geq r} (t_i + 1) \\ &\Leftrightarrow |I^-|s \geq \sum_{i \in I^-} t_i + |\{i \in I^-, r_i \geq r\}| \geq \sum_{i \in I^-} t_i + h_2. \end{aligned}$$

Finalement, on obtient l'inégalité :

$$c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i \in \{1, \dots, p\}} t_i - k + p = f(\mathcal{P}, q).$$

2.2.1.1.2 $r \neq 0$ et $p - k - r < 0$

Dans ce cas tous les termes de la suite $(r_i - r)$ sont strictement négatifs et donc $\forall i \in I^+$, $r_i'' = r_i - r + q - 1$. La configuration est donc la même qu'au point 2 de 2.2.1.1.1 ; la suite des r_i'' est ordonnée et s'écrit :

$$\underbrace{-r + q - 1, \dots, -r + q - 1}_{\leq k - 1 \text{ fois}}, \beta_1 - r + q - 1, \beta_2 - r + q - 1, \dots, \beta_l - r + q - 1.$$

De la même façon que précédemment, on obtient

$$\min_{i \in I^+} (r_i'' - i) \geq -r + q - k.$$

On applique l'hypothèse de récurrence pour calculer $c_{\min}(\mathcal{P}'', \leq q-1)$:

$$c_{\min}(\mathcal{P}'', \leq q-1) = \sum_{i \in I^+} (t_i - s - 1) + \min(0, \min_{i \in I^+} (r_i'' - i)) + |I^+|$$

et donc

$$c_{\min}(\mathcal{P}'', \leq q-1) \geq \sum_{i \in I^+} (t_i - s - 1) + \min(0, -r + q - k) + |I^+|.$$

Or d'après (\star) et (\diamond) , on sait que :

$$c(\mathcal{P}, \mathcal{Q}) \geq c_{\min}(\mathcal{P}'', \leq q-1) + |P_{i_0} \cap Q_{j_0}|.$$

Ainsi, puisque $|P_{i_0} \cap Q_{j_0}| = t_{i_0} + (q-1)s + r$,

$$c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i \in I^+} (t_i - s - 1) + \min(0, -r + q - k) + |I^+| + t_{i_0} + (q-1)s + r$$

$$\begin{aligned} \Leftrightarrow c(\mathcal{P}, \mathcal{Q}) &\geq \sum_{i \in I^+ \cup i_0} t_i - \sum_{i \in I^+} s - |I^+| + \min(0, -r + q - k) + |I^+| + (q - 1)s + r \\ &\Leftrightarrow c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i \in I^+ \cup i_0} t_i - |I^+|s + \min(0, -r + q - k) + (q - 1)s + r. \end{aligned}$$

Avec $|I^+| + |I^-| = p - 1$, il vient

$$c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i \in I^+ \cup i_0} t_i + |I^-|s + \min(0, -r + q - k) + (q - p)s + r.$$

Or comme nous l'avons montré précédemment, pour tout i tel que $n_i'' < 0$, c'est-à-dire pour tout $i \in I^-$, $s \geq t_i$. Ainsi $|I^-|s = \sum_{i \in I^-} s \geq \sum_{i \in I^-} t_i$, et on obtient donc

$$c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i=1}^p t_i + \min(0, -r + q - k) + (q - p)s + r.$$

De plus puisque $q \geq p$, on a donc

$$c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i=1}^p t_i + \min(0, -r + q - k) + r.$$

Si $r \leq q - k$, on obtient $c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i=1}^p t_i + r$ et puisque dans ce cas $r > p - k$, il vient

$$c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i=1}^p t_i + p - k.$$

Si $r > q - k$, on obtient

$$c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i=1}^p t_i - r + q - k + r \geq \sum_{i=1}^p t_i + p - k.$$

Finalement,

$$c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i=1}^p t_i + p - k = f(\mathcal{P}, q).$$

2.2.1.1.3 $r = 0$

La suite $(r_i - r)_{i \in I^+}$, est donc en fait la suite $(r_i)_{i \in I^+}$ et s'écrit :

$$\underbrace{0, 0, \dots, 0}_{\leq k-1 \text{ fois}}, \beta_1, \beta_2, \dots, \beta_l.$$

$\leq k - 1$ fois

Or, si $r_i \neq q - 1$,

$$\begin{cases} t_i'' = t_i - s \\ r_i'' = r_i \end{cases}$$

si $r_i = q - 1$,

$$\begin{cases} t_i'' = t_i - s + 1 \\ r_i'' = 0. \end{cases}$$

On note m le nombre d'indices $i \in I^+$ tels que $r_i = q - 1$. Ces indices correspondent aux m derniers éléments de la suite des r_i puisque cette suite est croissante et majorée par $q - 1$. La suite $(r''_i)_{i \in I^+}$ ordonnée, notée $(u_i)_{i \in I^+}$, s'écrit donc :

$$\underbrace{0, 0, \dots, 0, 0, 0, \dots, 0}_{m \text{ fois}}, \beta_1, \beta_2, \dots, \beta_{l-m}.$$

$\leq k - 1$ fois

Ainsi le minimum des $r''_i - i$ correspond au plus grand indice \hat{i} tel que $r''_{\hat{i}} = 0$. Puisque $\hat{i} \leq m + k - 1$, $\min_{i \in I^+}(r''_i - i) = r''_{\hat{i}} - \hat{i} = -\hat{i} \geq 1 - k - m$. On applique l'hypothèse de récurrence pour calculer $c_{\min}(\mathcal{P}'', \leq q - 1)$:

$$c_{\min}(\mathcal{P}'', \leq q - 1) = \sum_{i \in I^+} t''_i + \min(0, \min_{i \in I^+}(r''_i - i)) + |I^+|.$$

Or

$$\begin{aligned} \sum_{i \in I^+} t''_i &= \sum_{i \in I^+, r_i \neq q-1} (t_i - s) + \sum_{i \in I^+, r_i = q-1} (t_i - s + 1) \\ &= \sum_{i \in I^+} (t_i - s) + m, \end{aligned}$$

et donc on a :

$$c_{\min}(\mathcal{P}'', \leq q - 1) = \sum_{i \in I^+} (t_i - s) + m + \min(0, \min_{i \in I^+}(r''_i - i)) + |I^+|.$$

En utilisant maintenant la minoration $\min_{i \in I^+}(r''_i - i) \geq 1 - k - m$ établie plus haut, et en remarquant que $1 - k - m \leq 0$ puisque $k \geq 1$ (cas 2.2.1), il vient :

$$c_{\min}(\mathcal{P}'', \leq q - 1) \geq \sum_{i \in I^+} (t_i - s) + m + 1 - k - m + |I^+| = \sum_{i \in I^+} (t_i - s) + 1 - k + |I^+|.$$

Comme précédemment, on revient à $c(\mathcal{P}, \mathcal{Q})$, et d'après (\star) et (\diamond) , on obtient :

$$\begin{aligned} c(\mathcal{P}, \mathcal{Q}) &\geq c_{\min}(\mathcal{P}'', \leq q - 1) + |P_{i_0} \cap Q_{j_0}| \\ &\Rightarrow c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i \in I^+} (t_i - s) + |I^+| + 1 - k + (q - 1)s + t_{i_0} \\ &\Leftrightarrow c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i \in I^+ \cup i_0} t_i - |I^+|s + |I^+| + 1 - k + (q - 1)s. \end{aligned}$$

Avec $|I^+| + |I^-| = p - 1$, il vient :

$$c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i \in I^+ \cup i_0} t_i + |I^-|s + |I^+| + 1 - k + (q - p)s.$$

Or nous avons montré que $\forall i \in I^-$, l'inégalité $r_i - r \geq 0$ entraîne $s \geq t_i + 1$. Ainsi $|I^-|s = \sum_{i \in I^-} s \geq \sum_{i \in I^-} (t_i + 1)$, et on obtient donc

$$c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i=1}^p t_i + |I^-| + |I^+| + 1 - k + (q - p)s.$$

Finalement, puisque $q \geq p$ et $I^+ + I^- = p - 1$, il vient :

$$c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i=1}^p t_i + p - k = f(\mathcal{P}, q).$$

2.2.1.2 $i_0 > k$

Tout d'abord, notons que dans ce cas $\epsilon_{i_0} = 1$ et donc

$$|P_{i_0} \cap Q_{j_0}| = \alpha + t_{i_0} + 1.$$

On rappelle (cf. début du cas 2.2.1) que $\forall i \in \{1, \dots, p\}$,

$$|P_i \cap Q_{j_0}| \leq \alpha + t_i + \epsilon_i.$$

Si $k < i < i_0$, cette inégalité est stricte car sinon on aurait choisi $i_0 = i$:
 $\forall i \in \{k + 1, \dots, i_0 - 1\}$,

$$|P_i \cap Q_{j_0}| < \alpha + t_i + \epsilon_i.$$

Or $\epsilon_i = 0$ si $1 \leq i \leq k$ et $\epsilon_i = 1$ si $k < i \leq p$. Finalement, on a donc :

$$\begin{aligned} - \text{ si } 1 \leq i < i_0, & \quad |P_i \cap Q_{j_0}| \leq \alpha + t_i \\ - \text{ si } i_0 \leq i \leq p, & \quad |P_i \cap Q_{j_0}| \leq \alpha + t_i + 1. \end{aligned}$$

En posant $\forall i \in \{1, \dots, p\}$, $\nu_i = 0$ si $1 \leq i < i_0$ et $\nu_i = 1$ si $i_0 \leq i \leq p$, il vient

$$\forall i \in \{1, \dots, p\}, |P_i \cap Q_{j_0}| \leq \alpha + t_i + \nu_i.$$

En remplaçant maintenant $|P_i \cap Q_{j_0}|$ par $n_i - n'_i$, on obtient :

$$\forall i \in \{1, \dots, p\}, n'_i \geq n_i - \alpha - t_i - \nu_i.$$

Après utilisation des égalités $\alpha = (q - 1)s + r$ et $n_i = qt_i + r_i$, il vient :

$$n'_i \geq (q - 1)(t_i - s) + r_i - r - \nu_i.$$

Pour tout $i \in \{1, \dots, p\}$, on définit alors n''_i par :

$$n''_i = (q - 1)(t_i - s) + r_i - r - \nu_i.$$

On définit comme dans le cas 2.2.1.1 les ensembles I^+ et I^- . On redéfinit h_1 et h_2 par $h_1 = |\{i \in I^-, r_i - \nu_i < r\}|$ et $h_2 = |\{i \in I^-, r_i - \nu_i \geq r\}|$. On note encore \mathcal{P}'' la pseudo-partition à $|I^+|$ classes de cardinaux n''_i pour $i \in I^+$. On pose à nouveau $n''_i = t''_i(q - 1) + r''_i$, avec $r''_i \in \{0, \dots, q - 2\}$. Alors :

- si $r_i - r - \nu_i < 0$,

$$\begin{cases} t''_i &= t_i - s - 1 \\ r''_i &= r_i - r - \nu_i + q - 1 \end{cases}$$

- si $0 \leq r_i - r - \nu_i < q - 1$

$$\begin{cases} t''_i &= t_i - s \\ r''_i &= r_i - r - \nu_i \end{cases}$$

- si $r_i - r - \nu_i = q - 1$

$$\begin{cases} t''_i &= t_i - s + 1 \\ r''_i &= 0. \end{cases}$$

Par définition, $\forall i \in \{1, \dots, p\}$, $n'_i \geq n''_i$, ce qui implique de nouveau, d'après le Lemme 4 :

$$c_{\min}(\mathcal{P}', \leq q-1) \geq c_{\min}(\mathcal{P}'', \leq q-1). \quad (\square)$$

Écrivons la suite $(r_i - r - \nu_i)_{i \in \{1, \dots, p\}, i \neq i_0}$:

$$\underbrace{-r, -r, \dots, -r}_{k \text{ fois}}, 1 - r, \dots, i_0 - 1 - k - r, i_0 + 1 - k - r - 1, \dots, p - k - r - 1.$$

On remarque que le changement de valeur de la suite (ν_i) entre i_0 et $i_0 + 1$ est compensé par l'absence du terme d'indice i_0 dans la suite, qui croît alors régulièrement (de 1 en 1) à partir de l'indice k . Cette suite restreinte à I^+ s'écrit donc :

$$\underbrace{-r, -r, \dots, -r}_{\leq k \text{ termes}}, \underbrace{\beta_1 - r, \beta_2 - r, \dots, \beta_l - r}_{\ell \leq p - k - 1 \text{ termes}}$$

pour des β_i appropriés avec $\beta_i \geq 0$, $\beta_{i+1} > \beta_i$ et $\beta_l \leq p - k - 1$.
On distingue trois cas.

2.2.1.2.1 $r \neq 0$ et $p - k - r - 1 \geq 0$

- Si $\beta_l - r \geq 0$, on note j le plus petit indice appartenant à $\{1, \dots, l\}$ vérifiant $\beta_j - r \geq 0$. Dans la suite $(r_i - r - \nu_i)_{i \in \{1, \dots, p\}, i \neq i_0}$, il y a $r + k - 1$ termes strictement négatifs et $p - r - k$ termes positifs ou nuls, ce qui donne, pour la suite restreinte à I^+ , $r + k - 1 - h_1$ termes strictement négatifs et $p - r - k - h_2$ termes positifs ou nuls.

Soit $(u_i)_{i \in I^+}$ la suite (r''_i) ordonnée. Cette suite s'écrit : \leq
 k termes

$$\underbrace{\beta_{j+1} - r, \beta_{j+2} - r, \dots, \beta_l - r}_{p - k - r - h_2 \text{ termes}}, \underbrace{-r + q - 1, \dots, -r + q - 1, \beta_1 - r + q - 1, \dots, \beta_j - r + q - 1}_{k + r - 1 - h_1 \text{ termes}}.$$

Le minimum des $u_i - i$ est donc atteint soit pour $i = 1$, soit pour le plus grand indice, noté \hat{i} , tel que $u_i = -r + q - 1$ ($\hat{i} \leq p - r - h_2$). Ainsi

$$\min_{i \in I^+} (u_i - i) = \min(\beta_{j+1} - r - 1, -r + q - 1 - \hat{i}) \geq \min(\beta_{j+1} - r - 1, q - p - 1 + h_2) \geq -1.$$

- Si $\beta_l - r < 0$, tous les termes de la suite $(r_i - r - \nu_i)_{i \in \{1, \dots, p\}, i \neq i_0}$ sont strictement négatifs, et donc $\forall i \in I^+$, $r''_i = r_i - r - \nu_i + q - 1$. La suite des r''_i est donc ordonnée comme celle des r_i et s'écrit :

$$\underbrace{-r + q - 1, \dots, -r + q - 1}_{\leq k \text{ termes}}, \beta_1 - r + q - 1, \beta_2 - r + q - 1, \dots, \beta_l - r + q - 1$$

Le minimum des $r''_i - i$, $i \in I^+$, est donc atteint pour le plus grand indice, noté \hat{i} , tel que $r''_{\hat{i}} = -r + q - 1$. On a $\hat{i} \leq k$ et donc

$$\min_{i \in I^+} (r''_i - i) = r''_{\hat{i}} - \hat{i} \geq -r + q - 1 - k.$$

Or nous avons supposé que $p - k - r - 1 \geq 0$, et donc puisque $p \leq q$, nous avons aussi :

$$\min_{i \in I^+} (u_i - i) \geq -1.$$

On peut maintenant calculer un minorant de $c_{\min}(\mathcal{P}'', \leq q - 1)$ en utilisant l'hypothèse de récurrence :

$$\begin{aligned} c_{\min}(\mathcal{P}'', \leq q - 1) &= \sum_{i \in I^+} t_i'' + \min(0, \min_{i \in I^+} (u_i - i)) + |I^+| \\ \Rightarrow c_{\min}(\mathcal{P}'', \leq q - 1) &\geq \sum_{i \in I^+, r_i - r - \nu_i \geq 0} (t_i - s) + \sum_{i \in I^+, r_i - r - \nu_i < 0} (t_i - s - 1) - 1 + |I^+| \\ \Leftrightarrow c_{\min}(\mathcal{P}'', \leq q - 1) &\geq \sum_{i \in I^+} (t_i - s) - |\{i \in I^+, r_i - r - \nu_i < 0\}| - 1 + |I^+| \\ \Leftrightarrow c_{\min}(\mathcal{P}'', \leq q - 1) &\geq \sum_{i \in I^+} t_i - |I^+|s - (r + k - 1 - h_1) - 1 + |I^+|. \end{aligned}$$

Or on sait d'après (\star) et (\square) que

$$c(\mathcal{P}, \mathcal{Q}) \geq c_{\min}(\mathcal{P}'', \leq q - 1) + |P_{i_0} \cap Q_{j_0}|,$$

et donc, puisque $|P_{i_0} \cap Q_{j_0}| = \alpha + t_{i_0} + \epsilon_{i_0} = (q - 1)s + r + t_{i_0} + 1$, il vient :

$$c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i \in I^+ \cup i_0} t_i - |I^+|s + (q - p)s + (p - 1)s - k + 1 + h_1 + |I^+|$$

et, puisque $p \leq q$,

$$c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i \in I^+ \cup i_0} t_i - |I^+|s + (p - 1)s - k + 1 + h_1 + |I^+|.$$

Avec $|I^+| + |I^-| = p - 1$, il vient :

$$\begin{aligned} c(\mathcal{P}, \mathcal{Q}) &\geq \sum_{i \in I^+ \cup \{i_0\}} t_i - |I^+|s + (|I^+| + |I^-|)s - k + 1 + h_1 + (p - 1 - |I^-|) \\ \Leftrightarrow c(\mathcal{P}, \mathcal{Q}) &\geq \sum_{i \in I^+ \cup \{i_0\}} t_i + |I^-|s - k + h_1 - |I^-| + p. \end{aligned}$$

En utilisant maintenant $|I^-| = h_1 + h_2$, on obtient :

$$c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i \in I^+ \cup \{i_0\}} t_i + |I^-|s - k - h_2 + p.$$

On cherche maintenant à minorer s . Revenons à la définition de n_i'' :

$$n_i'' = (q - 1)(t_i - s) + r_i - r - \nu_i.$$

Or puisque $r_i \in \{0, \dots, q - 1\}$, $r \in \{0, \dots, q - 2\}$ et $\nu_i \in \{0, 1\}$, on a $r_i - r - \nu_i \geq 1 - q$ et donc

$$n_i'' \geq (q - 1)(t_i - s) - q + 1 = (q - 1)(t_i - s - 1).$$

Donc si $n_i'' < 0$, on a $t_i - s - 1 \leq -1$ ce qui équivaut à $s \geq t_i$. Si de plus $r_i - r - \nu_i \geq 0$, $n_i'' = (q-1)(t_i - s) + r_i - r - \nu_i < 0$ implique que $t_i - s < 0$ et donc $s \geq t_i + 1$. Ainsi

$$\begin{aligned} |I^-|s &= \sum_{i \in I^-} s = \sum_{i \in I^-, r_i - \nu_i < r} s + \sum_{i \in I^-, r_i - \nu_i \geq r} s \geq \sum_{i \in I^-, r_i - \nu_i < r} t_i + \sum_{i \in I^-, r_i - \nu_i \geq r} (t_i + 1) \\ &\Rightarrow |I^-|s \geq \sum_{i \in I^-} t_i + |\{i \in I^-, r_i - \nu_i \geq r\}| \geq \sum_{i \in I^-} t_i + h_4. \end{aligned}$$

Finalement, on obtient l'inégalité :

$$c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i=1}^p t_i - k + p = f(\mathcal{P}, q).$$

2.2.1.2.2 $r \neq 0$ et $p - k - r - 1 < 0$

Dans ce cas, $\forall i \in I^+$, $r_i - r - \nu_i < 0$, et donc $\forall i \in I^+$, $t_i'' = t_i - s - 1$ et $r_i'' = r_i - r - \nu_i + q - 1$. La suite des r_i'' est alors ordonnée et s'écrit :

$$\underbrace{-r + q - 1, \dots, -r + q - 1}_{\leq k \text{ fois}}, \beta_1 - r + q - 1, \beta_2 - r + q - 1, \dots, \beta_l - r + q - 1$$

Le minimum des $r_i'' - i$, $i \in I^+$, est donc atteint pour le plus grand indice, noté \hat{i} , tel que $r_{\hat{i}}'' = -r + q - 1$. On a $\hat{i} \leq k$ et donc

$$\min_{i \in I^+} (r_i'' - i) = r_{\hat{i}}'' - \hat{i} \geq -r + q - 1 - k.$$

On applique maintenant l'hypothèse de récurrence pour calculer $c_{\min}(\mathcal{P}'', \leq q-1)$:

$$\begin{aligned} c_{\min}(\mathcal{P}'', \leq q-1) &= \sum_{i \in I^+} t_i'' + \min(0, \min_{i \in I^+} (r_i'' - i)) + |I^+| \\ &\Rightarrow c_{\min}(\mathcal{P}'', \leq q-1) \geq \sum_{i \in I^+} (t_i - s - 1) + \min(0, -r + q - 1 - k) + |I^+|. \end{aligned}$$

Or on sait d'après (\star) et (\square) que

$$c(\mathcal{P}, \mathcal{Q}) \geq c_{\min}(\mathcal{P}'', \leq q-1) + |P_{i_0} \cap Q_{j_0}|,$$

et donc il vient :

$$\begin{aligned} c(\mathcal{P}, \mathcal{Q}) &\geq \sum_{i \in I^+} (t_i - s - 1) + \min(0, -r + q - 1 - k) + |I^+| + (q-1)s + r + t_{i_0} + 1 \\ &\Leftrightarrow c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i \in I^+ \cup i_0} t_i - \sum_{i \in I^+} s + \min(0, -r + q - 1 - k) + (q-1)s + r + 1. \end{aligned}$$

En utilisant $|I^+| + |I^-| = p - 1$, on obtient :

$$\begin{aligned} c(\mathcal{P}, \mathcal{Q}) &\geq \sum_{i \in I^+ \cup i_0} t_i + \sum_{i \in I^-} s + \min(0, -r + q - 1 - k) + (q-p)s + r + 1 \\ &\Rightarrow c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i \in I^+ \cup i_0} t_i + \sum_{i \in I^-} s + \min(0, -r + q - 1 - k) + r + 1 \end{aligned}$$

puisque $q \geq p$. De plus puisque $\forall i \in I^-$, on peut montrer comme dans le cas 2.2.1.2.1 que $s \geq t_i$, et donc :

$$c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i=1}^p t_i + \min(0, -r + q - 1 - k) + r + 1.$$

Finalement :

- si $-r + q - 1 - k \geq 0$, on a alors :

$$c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i=1}^p t_i + r + 1 \geq \sum_{i=1}^p t_i + p - k$$

car on a $r + 1 > p - k$ par hypothèse du cas 2.2.1.2.2 ;

- sinon on obtient :

$$c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i=1}^p t_i + r + 1 - r + q - 1 - k = \sum_{i=1}^p t_i + q - k \geq \sum_{i=1}^p t_i + p - k.$$

Ainsi dans les deux cas,

$$c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i=1}^p t_i + p - k = f(\mathcal{P}, q).$$

2.2.1.2.3 $r = 0$

La suite $(r_i - r - \nu_i)_{i \in I^+}$, est donc en fait la suite $(r_i - \nu_i)_{i \in I^+}$ et s'écrit :

$$\underbrace{0, 0, \dots, 0}_{\leq k \text{ fois}}, \beta_1, \beta_2, \dots, \beta_l.$$

Or, si $r_i - \nu_i \neq q - 1$,

$$\begin{cases} t_i'' &= t_i - s \\ r_i'' &= r_i - \nu_i. \end{cases}$$

si $r_i - \nu_i = q - 1$,

$$\begin{cases} t_i'' &= t_i - s + 1 \\ r_i'' &= 0. \end{cases}$$

On note m le nombre d'indices $i \in I^+$ tels que $r_i - \nu_i = q - 1$. Ces indices correspondent aux m derniers éléments de la suite des $r_i - r - \nu_i$ puisque cette suite est croissante. La suite $(r_i'')_{i \in I^+}$ ordonnée, notée $(u_i)_{i \in I^+}$, s'écrit donc :

$$\underbrace{0, 0, \dots, 0}_{m \text{ fois}}, \underbrace{0, 0, \dots, 0}_{\leq k \text{ fois}}, \beta_1, \beta_2, \dots, \beta_{l-m}.$$

Ainsi le minimum des $r_i'' - i$ correspond au plus grand indice \hat{i} tel que $r_i'' = 0$. Puisque $\hat{i} \leq m + k$, $\min_{i \in I^+} (r_i'' - i) = r_{\hat{i}}'' - \hat{i} = -\hat{i} \geq -k - m$. On applique l'hypothèse de récurrence pour calculer $c_{\min}(\mathcal{P}'', \leq q - 1)$:

$$c_{\min}(\mathcal{P}'', \leq q - 1) = \sum_{i \in I^+} t_i'' + \min(0, \min_{i \in I^+} (r_i'' - i)) + |I^+|.$$

Or

$$\begin{aligned} \sum_{i \in I^+} t_i'' &= \sum_{i \in I^+, r_i - r - \nu_i \neq q-1} (t_i - s) + \sum_{i \in I^+, r_i - r - \nu_i = q-1} (t_i - s + 1) \\ &= \sum_{i \in I^+} (t_i - s) + m, \end{aligned}$$

et donc on a :

$$c_{\min}(\mathcal{P}'', \leq q-1) = \sum_{i \in I^+} (t_i - s) + m + \min(0, \min_{i \in I^+} (r_i'' - i)) + |I^+|.$$

En utilisant maintenant la minoration précédente, et puisque $\min_{i \in I^+} (r_i'' - i) < 0$, il vient :

$$c_{\min}(\mathcal{P}'', \leq q-1) \geq \sum_{i \in I^+} (t_i - s) + m - k - m + |I^+| = \sum_{i \in I^+} (t_i - s) - k + |I^+|.$$

Comme précédemment, on revient à $c(\mathcal{P}, \mathcal{Q})$, et d'après (\star) et (\square) , on obtient :

$$c(\mathcal{P}, \mathcal{Q}) \geq c_{\min}(\mathcal{P}'', \leq q-1) + |P_{i_0} \cap Q_{j_0}|$$

$$\Rightarrow c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i \in I^+} (t_i - s) + |I^+| - k + (q-1)s + t_{i_0} + 1$$

$$\Leftrightarrow c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i \in I^+ \cup i_0} t_i - |I^+|s + |I^+| - k + (q-1)s + 1.$$

Avec $|I^+| + |I^-| = p-1$, il vient :

$$c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i \in I^+ \cup i_0} t_i + |I^-|s + |I^+| - k + (q-p)s + 1.$$

On peut montrer, de la même façon que dans le cas 2.2.1.2.1, que l'inégalité $r_i - \nu_i \geq 0$, pour tout $i \in I^-$, entraîne $s \geq t_i + 1$.

Ainsi $|I^-|s = \sum_{i \in I^-} s \geq \sum_{i \in I^-} (t_i + 1)$, et on obtient donc

$$c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i=1}^p t_i + |I^-| + |I^+| - k + (q-p)s + 1.$$

Finalement, puisque $q \geq p$ et $I^+ + I^- = p-1$, il vient :

$$c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i=1}^p t_i + p - k = f(\mathcal{P}, q).$$

- **cas 2.2.2** : $k = 0$ Supposons maintenant que $k = 0$. Les p classes de la partition \mathcal{P}

possèdent alors des restes distincts et non nuls ; on a $r_p = p$ et puisque r_p est un reste de la division entière par q , ceci implique que $p < q$. On considère alors la pseudo-partition $\tilde{\mathcal{P}}$ à $p+1$ classes ($p+1 \leq q$) composée des p classes de \mathcal{P} et d'une classe vide. Il est clair que $c_{\min}(\mathcal{P}, \leq q) = c_{\min}(\tilde{\mathcal{P}}, \leq q)$ puisque l'ajout d'une classe vide ne change pas le calcul de la concordance.

La suite des restes ordonnés (\tilde{r}_i) de $\tilde{\mathcal{P}}$ est : $0, r_1, r_2, \dots, r_p$ et donc

$$\min(0, \min_{1 \leq i \leq p+1} (\tilde{r}_i - i)) = \min(0, \min_{1 \leq i \leq p} (r_i - i)) - 1.$$

Ainsi :

$$f(\tilde{\mathcal{P}}, q) = \sum_{i=1}^p t_i + 0 + \min(0, \min_{1 \leq i \leq p+1} (\tilde{r}_i - i)) + p + 1 = f(\mathcal{P}, q).$$

Finalemnt montrer que $c_{\min}(\mathcal{P}, \leq q) = f(\mathcal{P}, q)$ revient à montrer l'égalité entre $c_{\min}(\tilde{\mathcal{P}}, \leq q)$ et $f(\tilde{\mathcal{P}}, q)$, ce qui correspond au cas $k \geq 1$, étudié précédemment (cas 2.2.1).

Nous avons donc montré que pour toute pseudo-partition \mathcal{Q} quelconque possédant q classes,

$$c(\mathcal{P}, \mathcal{Q}) \geq \sum_{i=1}^p t_i + p - k = f(\mathcal{P}, q),$$

ce qui implique que

$$c_{\min}(\mathcal{P}, \leq q) \geq f(\mathcal{P}, q)$$

et donc d'après le Lemme 2,

$$c_{\min}(\mathcal{P}, \leq q) = f(\mathcal{P}, q).$$

On a donc montré que si le Théorème 3 est vrai pour toute partition sur un ensemble à moins de n éléments, alors il est vrai pour toute partition sur n éléments, ce qui achève la démonstration par récurrence du Théorème 3.

□