Optimisation de la transmission d'information sur un canal quantique via le calcul par intervalles

Nicolas DELANOUE, François CHAPEAU-BLONDEAU,

Laboratoire Angevin de Recherche en Ingénierie des Systèmes (LARIS),
Université d'Angers, 62 avenue Notre Dame du Lac, 49000 Angers, France.
nicolas.delanoue@univ-angers.fr, f.chapeau@univ-angers.fr

Résumé – La transmission d'information sur un canal quantique de communication soulève des problèmes d'optimisation spécifiques, non encore universellement résolus. Nous considérons ici l'application de méthodes d'optimisation globale basées sur le calcul par intervalles, pour la maximisation de critères informationnels caractérisant la performance d'un canal quantique de communication, en insistant spécialement sur leur apport et potentiel spécifiques.

Abstract – Transmission of information over a quantum communication channel raises specific optimization problems, not yet universally solved. Here we consider the application of global optimization methods based on interval computation, to maximize informational criteria characterizing the performance of a quantum communication channel, with special emphasis on their specific contribution and potential.

1 Introduction

La théorie quantique de l'information s'intéresse, entre autre, à la communication efficace d'information sur un canal quantique [1, 2]. En pratique, on rencontre ce genre de situations lorsque les entités supportant l'information sont des objets élémentaires, comme des photons individuels sur une fibre optique, qui doivent être traités conformément à leur nature quantique. La théorie quantique de l'information présente une grande parenté avec la classique, mais également des particularités qui lui sont propres. Communément, en présence d'un canal de transmission donné, on souhaite optimiser ses conditions d'utilisation, comme par exemple les signaux appliqués en entrée ou leur traitement en sortie, afin de maximiser un critère de performance approprié. En quantique, on retrouve le critère d'information mutuelle entrée-sortie, mais il se trouve complété par d'autres, différemment pertinents, comme l'information de Holevo, ou l'information accessible. Aussi, en quantique, les variables à ajuster dans l'optimisation sont de nature spécifique, typiquement des opérateurs densité pour les signaux et des opérateurs de projection pour la mesure. On rencontre ainsi en information quantique des problèmes d'optimisation spécifiques, et qui ne sont pas encore universellement résolus. Parmi les problèmes pour lesquels il existe une approche générale, on peut citer [3] où les auteurs conçoivent un détecteur quantique optimal en minimisant la probabilité d'erreur de détection. La conception optimale de [3] optimise ainsi un critère linéaire, et se réduit à la résolution d'un problème de programmation semi-définie. Plus généralement, en présence de critères non linéaires, on dispose de méthodes d'optimisation présentant des apports spécifiques qui sont les approches par intervalles. Elles permettent notamment d'obtenir des optimums globaux garantis, ou bien des caractérisations ensemblistes des solutions admissibles satisfaisant un critère. Nous abordons et exposons ici, pour la première fois à notre connaissance, l'application d'approches intervalistes sur des problèmes d'optimisation typiques rencontrés en théorie quantique de l'information, en insistant notamment sur leur apport et potentiel spécifiques. Ce faisant, ceci nous permet de proposer un exposé synthétique des méthodes intervalistes pour l'optimisation, ainsi qu'une formulation synthétique de problèmes d'optimisation clés de l'information quantique.

2 Optimisation globale via le calcul par intervalles

Un intervalle $[\underline{x}, \overline{x}]$, aussi noté $[x]$, est un ensemble de la forme $\{x \in \mathbb{R}^n, \underline{x} \leq x \leq \overline{x}\}$ avec $\underline{x}$ et $\overline{x}$ deux éléments de $\mathbb{R}^n$. Ici, la relation $\leq$ est entendue composante par composante. Les intervalles considérés ici sont donc nécessairement bornés. L'ensemble des intervalles est généralement désigné par $\mathbb{IR}^n$.

Définition 2.1. *Une fonction* $[f] : \mathbb{IR}^n \to \mathbb{IR}^m$ *est une fonction d'inclusion pour* $f : \mathbb{R}^n \to \mathbb{R}^m$ *si* $\forall [x] \in \mathbb{IR}^n, f([x]) \subset [f]([x])$ *(où* $f([x]) = \{f(x) | x \in [x]\}$*).*

C'est-à-dire qu'une fonction d'inclusion $[f]$ associe à chaque intervalle $[x]$ de $\mathbb{IR}^n$ un intervalle $[f]([x])$ contenant l'image directe de $[x]$ par f. Cette propriété est illustrée par la Fig. 1.

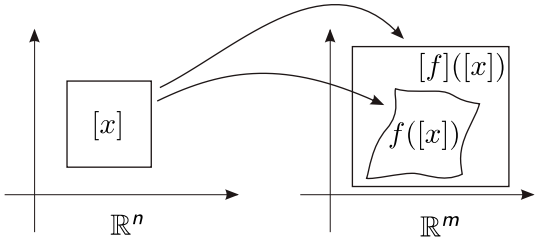

FIGURE 1 – Une fonction d'inclusion $[f]$ pour f.

L'arithmétique des intervalles [7] fournit une méthode effective pour construire des fonctions d'inclusion. La Réf. [8]

prouve qu'il est toujours possible d'obtenir une fonction d'inclusion $[f]$ lorsque f est définie par une expression arithmétique. Cette possibilité de borner l'image directe d'un intervalle $[x]$ par f permet par exemple de certifier des propriétés portant sur une quantité non dénombrable de nombres réels. En effet, si $0 \notin [f]([x])$, alors $\forall x \in [x], f(x) \neq 0$. Depuis l'introduction de l'arithmétique des intervalles par [6] et [7], de nombreux algorithmes ont été développés dans différents domaines : en optimisation globale ou bien encore pour les systèmes dynamiques non linéaires. Comme le calcul par intervalles produit des valeurs dites garanties, des algorithmes s'appuyant sur ces techniques peuvent être utilisés pour prouver des affirmations mathématiques comme la conjecture de Kepler ou bien l'existence de l'attracteur de Lorentz. Dans cet article, le calcul par intervalles est considéré pour l'optimisation globale de critères informationnels caractérisant la performance d'un canal quantique de communication.

2.1 Bornes garanties via le calcul par intervalles

Définition 2.2 (Arithmétique d'intervalles)**.** *Soit* $[x] = [\underline{x}, \overline{x}]$ *et* $[y] = [\underline{y}, \overline{y}]$ *deux intervalles,* i.e. $[x], [y] \in \mathbb{IR}$ *. Comme dans [7], l'arithmétique d'intervalles est définie par les quatre opérations suivantes :*

$$\begin{array}{rcl} [x] + [y] & = & [\underline{x} + \underline{y}, \overline{x} + \overline{y}], \\ [x] - [y] & = & [\underline{x} - \overline{y}, \overline{x} - \underline{y}], \\ [x] \times [y] & = & [\min\{\underline{x}\underline{y}, \underline{x}\overline{y}, \overline{x}\underline{y}, \overline{x}\overline{y}\} \ , \ \max\{\underline{x}\underline{y}, \underline{x}\overline{y}, \overline{x}\underline{y}, \overline{x}\overline{y}\}], \\ [x] \div [y] & = & [x] \times \left[\frac{1}{\overline{y}}, \frac{1}{\underline{y}}\right] \ , \ si \ \underline{y}\overline{y} > 0. \end{array} \tag{1}$$

Les quatre opérations (1) sont des fonctions d'inclusion pour les opérations classiques de l'arithmétique définie sur les nombres réels. De plus, la notion de fonctions d'inclusion est stable par composition.

Exemple 2.1. Soit $f : \mathbb{R}^2 \to \mathbb{R}$ avec $f(x, y) = 1 - 2y + x^2$. La fonction $[f] : \mathbb{IR}^2 \to \mathbb{IR}$ définie par $[f]([x], [y]) = [1, 1] - [2, 2] \times [\underline{y}, \overline{y}] + [\underline{x}, \overline{x}]^2$, est une fonction d'inclusion pour f.

2.2 Optimisation globale

Le calcul par intervalles permet donc de calculer des bornes garanties pour l'image d'un intervalle pour une fonction réelle donnée [8]. En combinant ces calculs de bornes avec une approche de type *séparation et évaluation* (branch and bound), on peut résoudre des problèmes d'optimisation comme :

$$f^* = \sup_{x \in G} f(x), \tag{2}$$

avec $G = \{x \in \mathbb{R}^n, g_i(x) \leq 0, i = 1 \dots q\}$. Sous certaines hypothèses [8], ce genre d'approche génère une suite d'intervalles $\{[f_i]\}_{i \in \mathbb{N}}$ convergente vers le singleton $\{f^*\}$ telle que $\forall i \in \mathbb{N}, f^* \in [f_i]$. La phase de séparation consiste à diviser le problème en un certain nombre de sous-problèmes qui ont chacun leur ensemble de solutions, de telle sorte que tous ces ensembles forment un recouvrement de l'ensemble admissible G. L'évaluation a pour but de déterminer des bornes inférieures et supérieures contenant l'ensemble des valeurs possibles d'un sous-problème donné. En combinant ces deux phases, il est possible d'écarter un ensemble ne contenant nécessairement pas de solution optimale. La proposition 2.1 [8], illustrée par la Fig. 2, formalise cette approche.

Proposition 2.1. *Soit* $\mathcal{G} = \{G_j\}_{j \in J}$ *un recouvrement de* G *et* $\{\overline{f}_j\}_{j \in J}$ *des nombres réels vérifiant* $\forall x \in G_j, f(x) \leq \overline{f}_j$*,* $\hat{x} \in G$ *un candidat et* x^* *une solution optimale de (2) alors* $\overline{f}_k \leq f(\hat{x}) \Rightarrow x^* \notin G_k$.

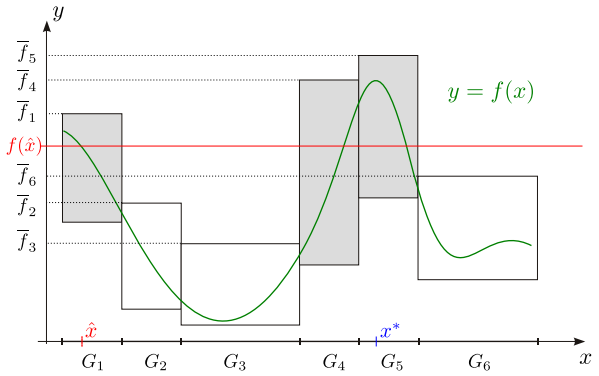

FIGURE 2 – Ici, $G = \cup_{j=1}^{j=6} G_j$, la solution x^* n'appartient ni à G_2, ni à G_3, ni à G_6.

L'algorithme basique de séparation et évaluation du calcul par intervalles est un algorithme récursif. Pour chaque élément G_j d'un recouvrement de l'ensemble des solutions optimales $\mathcal{G}$ donné, on calcule un candidat $\hat{x}_j$ et une borne supérieure $\overline{f}_j$ puis on nomme $\hat{x}$ le meilleur candidat. À partir de $f(\hat{x})$, tous les ensembles G_k du recouvrement $\mathcal{G}$ vérifiant $\overline{f}_k \leq f(\hat{x})$ sont écartés de l'espace de recherche. Formellement, on construit le nouveau recouvrement $\mathcal{G}'$ via $\mathcal{G}' = \mathcal{G} - \{G_k \mid \overline{f}_k \leq f(\hat{x})\}$. Puis, chaque ensemble de $\mathcal{G}'$ est lui-même fractionné afin d'obtenir un recouvrement plus fin et la procédure est récursivement appelée jusqu'à ce qu'un critère d'arrêt soit satisfait. Ce critère d'arrêt peut porter sur la taille des ensembles du recouvrement ou bien sur la précision de $f(\hat{x})$ relativement aux autres grandeurs $\underline{f}_j$ et $\overline{f}_j$. Cette procédure est initialement appelée avec un recouvrement de l'ensemble admissible G. Cette procédure jouit de nombreuses propriétés. À chaque appel récursif, le recouvrement contient nécessairement la solution optimale. De plus, sous des hypothèses de compacité de l'ensemble G et de continuité de la fonction d'inclusion $[f]$, cette méthode converge de façon garantie vers l'optimum global. En arrêtant la procédure à un instant donné, il est possible de fournir une solution admissible $\hat{x}$ et de mesurer sa qualité. En effet, l'écart maximum entre $f(\hat{x})$ et l'un des $\overline{f}_i$ donne la potentielle marge d'amélioration encore disponible. On parle alors de *gap* potentiel. Par ailleurs, avec des techniques d'inversion ensembliste [4, 5], il est aussi possible de caractériser l'ensemble de solutions admissibles étant au dessus d'un certain seuil. Cette technique est présentée dans la section suivante. Cette caractérisation ensembliste peut aussi être employée pour décrire le recouvrement $\{G_j\}_{j \in J}$ de la proposition 2.1 dans un contexte

d'optimisation sous contraintes.

2.3 Résolution ensembliste

Les techniques d'inversion ensembliste permettent de décrire géométriquement un ensemble G défini par des inégalités [4]. On pourra par exemple caractériser l'ensemble de vecteurs $x \in \mathbb{R}^m$ vérifiant l'inégalité $g(x) \geq 0$. Comme la procédure d'optimisation de la section précédente, l'algorithme d'inversion ensembliste peut être présenté récursivement avec comme entrée un recouvrement $\mathcal{G}$. Tout d'abord, on suppose qu'on dispose d'une fonction d'inclusion $[g]$ pour la fonction g, puis pour chaque élément G_j du recouvrement $\mathcal{G}$, on calcule des bornes $\underline{g}_j$ et $\overline{g}_j$. Plusieurs cas incompatibles peuvent se présenter : i) $\underline{g}_j \geq 0$, alors $G_j \subset S$, ii) $\overline{g}_j < 0$, alors $G_j \cap S = \emptyset$, iii) autrement, *i.e.* $0 \in]\underline{g}_j, \overline{g}_j]$ et on ne peut rien conclure. Dans les deux premiers cas, le statut des éléments de G_j est clair et cette information est stockée. La procédure est alors récursivement appelée avec comme entrée un recouvrement, plus fin, obtenu en fractionnant l'ensemble des G_j vérifiant la condition iii).

Exemple 2.2. On considère ici l'ensemble des matrices réelles symétriques $\mathsf{E} = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$ vérifiant les contraintes $\mathsf{E} \succeq 0$ et $\mathrm{Id} - \mathsf{E} \succeq 0$. Ces contraintes de semi-définie positivité peuvent se réécrire sous forme d'inégalités suivantes : $ad - b^2 \geq 0$ et $(1-a)(1-d) - b^2 \geq 0$ et $a \geq 0$ et $d \geq 0$ et $1 - a \geq 0$ et $1 - d \geq 0$. La Fig. 3 donne le résultat d'une inversion ensembliste de cet ensemble dans le cas où $d = 0,7$.

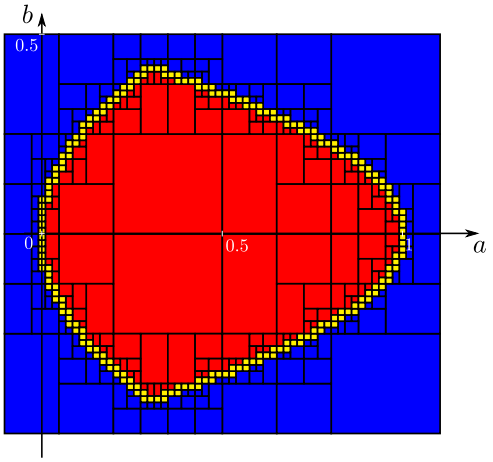

FIGURE 3 – Caractérisation ensembliste dans le plan (a, b) des matrices réelles E vérifiant $\mathsf{E}^\top = \mathsf{E}$, $\mathsf{E} \succeq 0$ et $\mathrm{Id} - \mathsf{E} \succeq 0$ avec $d = 0,7$. Les intervalles rouges, bleus et jaunes satisfont respectivement les conditions i), ii) et iii).

3 Canal quantique

Nous présentons maintenant des problèmes d'optimisation issus de l'information quantique. Une source d'information classique émet un symbole discret X prenant les valeurs x_j avec les probabilités a priori $p_j = \Pr\{X = x_j\}$, pour $j = 1$ à J. En vue d'une communication sur un canal quantique [1, 2], chaque symbole x_j est encodé par un état quantique représenté par l'opérateur densité $\rho_j \in \mathcal{L}(\mathcal{H}_N)$ sur l'espace de Hilbert complexe $\mathcal{H}_N$ de dimension N. Le canal quantique réalise une transformation d'un opérateur densité d'entrée ρ en un opérateur densité de sortie $\rho' = \mathcal{N}(\rho)$ qui peut s'écrire

$$\rho' = \mathcal{N}(\rho) = \sum_{\ell} \Lambda_\ell \rho \Lambda_\ell^\dagger \,, \tag{3}$$

où les opérateurs de Kraus $\{\Lambda_\ell\}$ de $\mathcal{L}(\mathcal{H}_N)$ définissent le modèle de canal (avec $\Lambda_\ell^\dagger$ l'opérateur adjoint de Λ_ℓ). En sortie du canal, on réalise une mesure quantique (généralisée ou "POVM") [1] définie par K opérateurs positifs $\mathsf{E}_k \in \mathcal{L}(\mathcal{H}_N)$, pour $k = 1$ à K, qui sont tenus de décomposer l'identité de $\mathcal{H}_N$, et qui établissent les probabilités conditionnelles du décodage en sortie fournies par les traces d'opérateurs,

$$\Pr\{Y = y_k | X = x_j\} = \mathrm{tr}(\rho'_j \mathsf{E}_k) \,, \tag{4}$$

avec $\rho'_j = \mathcal{N}(\rho_j)$. La transmission d'information $X \to Y$ par l'intermédiaire du canal quantique établissant les probabilités conditionnelles entrée-sortie de l'Éq. (4), peut alors être caractérisée par l'information mutuelle classique

$$I(X;Y) = H(Y) - H(Y|X) \,, \tag{5}$$

où $H(\cdot)$ est l'entropie de Shannon. Différents problèmes d'optimisation intéressants résultent alors. Pour un ensemble d'encodage $\{p_j, \rho_j\}$ fixé en entrée, il est utile de rechercher l'ensemble d'opérateurs de mesure $\{\mathsf{E}_k\}$ en sortie qui maximisent l'information mutuelle $I(X;Y)$ de l'Éq. (5) et définissent [1, 2] l'information accessible (sic)

$$I_{\mathrm{acc}}(X;Y) = \max_{\{\mathsf{E}_k\}} I(X;Y) \,. \tag{6}$$

Il ne s'agit pas d'un problème d'optimisation universellement résolu. En particulier, le cardinal K de l'ensemble solution est a priori inconnu, bien que majoré par N^2, et que les E_k optimaux peuvent être recherchés comme proportionnels à des projecteurs de rang 1. Une borne supérieure est par ailleurs connue [1, 2] pour l'information accessible $I_{\mathrm{acc}}(X;Y)$, constituée par l'information de Holevo

$$\chi\big(\{p_j, \rho'_j\}\big) = S\Big(\sum_{j=1}^{J} p_j \, \rho'_j\Big) - \sum_{j=1}^{J} p_j \, S(\rho'_j) \,, \tag{7}$$

où $S(\cdot)$ est l'entropie de von Neumann. Une autre optimisation envisageable [1, 9] consiste alors à rechercher l'ensemble d'encodage $\{p_j, \rho_j\}$ en entrée accomplissant la maximisation

$$\chi_{\max} = \max_{\{p_j, \rho_j\}} \chi\big(\{p_j, \rho'_j\}\big) \,. \tag{8}$$

Là aussi, il ne s'agit pas d'un problème d'optimisation universellement résolu. En particulier, le cardinal J de l'ensemble solution est a priori inconnu, bien que majoré par N^2, et que les ρ_j optimaux peuvent être recherchés comme des états purs, *i.e.* de la forme $\rho_j = |\psi_j\rangle \langle\psi_j|$ avec $|\psi_j\rangle \in \mathcal{H}_N$. D'autres problèmes d'optimisation non triviaux peuvent être envisagés, notamment si l'on impose d'autres contraintes sur les ensembles d'encodage $\{p_j, \rho_j\}$ ou de mesure $\{\mathsf{E}_k\}$.

4 Un détecteur quantique optimal

La conception d'un détecteur quantique optimal va consister ici à résoudre le problème d'optimisation décrit par l'Éq. (6). On considère l'instance suivante : on code en entrée du canal les deux symboles x_1 et x_2 avec les opérateurs densité $\rho_1 = |\psi_1\rangle\langle\psi_1|$, $\rho_2 = |\psi_2\rangle\langle\psi_2|$ et les deux états purs non orthogonaux $|\psi_1\rangle = \frac{1}{4}|0\rangle + \frac{\sqrt{15}}{4}|1\rangle$, $|\psi_2\rangle = \frac{\sqrt{2}}{2}|0\rangle + \frac{\sqrt{2}}{2}|1\rangle$, de probabilités a priori $p_1 = 0.1$, $p_2 = 0.9$. Le canal de transmission de l'Éq. (3) est ici pris comme l'identité. Il s'agit donc de déterminer la mesure quantique optimale permettant d'extraire le maximum d'information de deux états quantiques d'encodage imposés. Pour cette instance, on cherche deux opérateurs de mesure optimaux E_1 et E_2 maximisant l'information mutuelle entrée-sortie. Les opérateurs de mesure E_1 et E_2 vérifient nécessairement $\mathsf{E}_1 + \mathsf{E}_2 = \mathrm{Id}$. Comme ρ_1 et ρ_2 sont à coefficients réels, on peut supposer que les opérateurs E_k aussi et poser $\mathsf{E}_1 = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$, le problème de maximisation de l'information mutuelle se formule donc comme un problème d'optimisation non linéaire : $\max_{(a,b,d)\in\mathbb{R}^3} I(a,b,d)$ sous les contraintes $ad - b^2 \geq 0, (1-a)(1-d) - b^2 \geq 0, a \geq 0, b \geq 0, 1-a \geq 0, 1-d \geq 0$. Ici $I(a,b,d)$ est l'information mutuelle donnée par l'Éq. (5) réécrite en fonction des variables à optimiser : a, b et d. Une optimisation globale garantie, comme présentée dans la section 2, donne comme solution les deux opérateurs de mesure suivants $\mathsf{E}_1 = \begin{bmatrix} 0.446 & -0.497 \\ -0.497 & 0.554 \end{bmatrix}$ et $\mathsf{E}_2 = \mathrm{Id} - \mathsf{E}_1$. La valeur optimale de l'information mutuelle est de $0,0957$ Sh pour cette solution. De plus, l'exécution de notre algorithme certifie que le gap entre cette valeur et la valeur maximale est inférieure à 10^{-4} (avec un temps de calcul inférieur à 20 secondes). Cette valeur optimale pour l'information mutuelle est bien évidemment inférieure à l'information de Holevo (7) qui est de 0.1622 Sh pour cette instance. Il est aussi possible de caractériser l'ensemble des opérateurs de mesure dépassant une valeur donnée de l'information mutuelle (voir Fig. 4). Finalement, la Fig. 4 illustre le fait que l'information mutuelle présente une sorte de plateau au voisinage de la solution optimale. Cette caractérisation nous renseigne sur la précision nécessaire lors de la réalisation matérielle du détecteur.

5 Conclusion et perspectives

L'exemple traité montre, pour la première fois à notre connaissance, l'applicabilité et des apports des approches intervalistes pour aborder les problèmes d'optimisation spécifiques formulés en information quantique. De nombreux développements restent envisageables, comme l'étude du comportement des solutions optimales en présence de bruit ou de décohérence quantique, ou bien l'utilisation d'états quantiques intriqués optimalement en vue d'améliorer la transmission d'information.

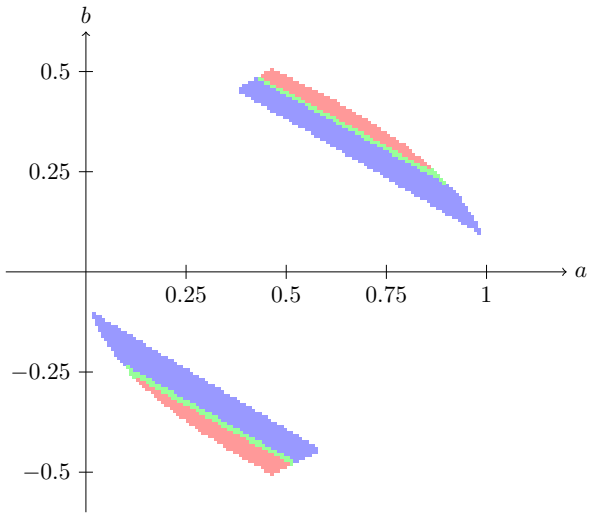

FIGURE 4 – Dans le plan (a, b), les trois zones (bleu, rouge et verte) caractérisent une section (à $d = 0.535$) de l'ensemble des solutions admissibles E_1 pour lesquelles l'information mutuelle est potentiellement supérieure à 0.05, 0.08 et 0.09, respectivement.

Références

[1] M. M. Wilde, "Quantum Information Theory", *Cambridge : Cambridge University Press*, 2017.

[2] C. H. Bennett, P. Shor, "Quantum information theory," *IEEE Transactions on Information Theory*, vol. 44, pp. 2724–2742, 1998.

[3] Y. C. Eldar, A. Megretski, G. C. Verghese, "Designing optimal quantum detectors via semidefinite programming", *IEEE Transactions on Information Theory*, vol. 49, pp. 1007–1012, 2003.

[4] L. Jaulin, M. Kieffer, O. Didrit, E. Walter, "Applied Interval Analysis with Examples in Parameter and State Estimation, Robust Control and Robotics", *Springer-Verlag*, 2001

[5] N. Delanoue, S. Lagrange. "A numerical approach to compute the topology of the apparent contour of a smooth mapping from $\mathbb{R}^2$ to $\mathbb{R}^2$". *Journal of Computational and Applied Mathematics*. vol. 271 pp. 267–284, 2014.

[6] T. Sunaga, "Theory of interval algebra and its application to numerical analysis", *Research Association of Applied Geometry (RAAG) Memoirs*, Vol. 2, pp. 29–46, 1958.

[7] R. E. Moore, "Interval Analysis", *Prentice-Hall, Englewood Cliffs*, NJ, 1966

[8] A. Neumaier, "Interval Methods for Systems of Equations" *Encyclopedia of Mathematics and its Applications* 37, Cambridge Univ. Press, Cambridge 1990

[9] F. Chapeau-Blondeau, "Optimization of quantum states for signaling across an arbitrary qubit noise channel with minimum-error detection," *IEEE Transactions on Information Theory*, vol. 61, pp. 4500–4510, 2015.

■