

## Sujet de thèse : **Maximum de vraisemblance et apprentissage informé pour l'imagerie dynamique en radioastronomie**

Encadrants : Isabelle Vin (MCF Université Paris-Saclay, [isabelle.vin@ens-paris-saclay.fr](mailto:isabelle.vin@ens-paris-saclay.fr)) et Pascal Larzabal (PU Université Paris-Saclay, [pascal.larzabal@universite-paris-saclay.fr](mailto:pascal.larzabal@universite-paris-saclay.fr))

École doctorale : STIC, pole A

Lieu : ENS Paris Saclay

Laboratoire : SATIE

Financement : Contrat École Doctorale – Programme blanc GS-SIS

Début : 1<sup>er</sup> octobre 2022

### Le contexte

Le “Square Kilometer Array” (SKA) [1-2] est un projet de radiotélescope géant, de surface collectrice équivalente à un kilomètre carré. Il est constitué de plusieurs réseaux interférométriques dans les longueurs d'onde métriques et centimétriques. Il est prévu de déployer SKA sur deux sites, l'un en Afrique du Sud (SKA-mid pour les hautes fréquences) et l'autre en Australie (SKA-low pour les basses fréquences). Le déploiement se déroulera en deux phases séparées dans le temps: La phase 1 (SKA1) a débuté en 2021 par la construction prévue pour 2030 d'environ 200 antennes paraboliques (15 m de diamètre, bande passante 350MHz-13GHz) en Afrique du Sud et 130 000 antennes phasées (antenne log périodique à double polarisation bande passante (50MHz-350MHz) travaillant aux basses fréquences dans l'ouest australien. Ce radiotélescope géant sera évolutif mais une mise en service partielle aura lieu dès 2024. Dans cette configuration, SKA1 représentera déjà un saut qualitatif immense par rapport aux instruments existants, et permettra des avancées décisives dans toutes les thématiques de l'astrophysique et de la physique modernes, comme la cosmologie, l'origine des champs magnétiques cosmiques, les ondes gravitationnelles, le milieu interstellaire, la formation des étoiles aux différentes époques de l'univers, les sursauts radio rapides... Depuis le début, SKA a été conçu pour accéder à un champ d'exploration exceptionnel au travers de la combinaison de caractéristiques (sensibilité, champ de vue, résolution, couverture en fréquence) qui seront d'un ordre de grandeur supérieur à celles des instruments radio à basses et moyennes fréquences existants. Une capacité sans égale d'obtenir des images très nettes avec des temps de pose assez courts permettra à ce nouvel observatoire de balayer le ciel rapidement et en profondeur. La Phase 2 est envisagée pour les années 2030+. Dans cette configuration finale, SKA2 sera l'instrument ultime de la radioastronomie basse-fréquence du 21<sup>ème</sup> siècle. Dès la phase1 SKA1 sera l'une des plus formidables machines jamais déployées par l'homme, et de loin la plus impressionnante en termes de débit de données et de puissance de calcul engagée.

### La problématique

L'imagerie radioastronomique du ciel est traditionnellement effectuée par des techniques d'interférométrie dont le pouvoir de *résolution* (finesse des détails de l'image) est directement lié à l'étendue du réseau d'antenne, c'est à dire la distance séparant ses antennes les plus éloignées [3-4]. Cela veut dire qu'avec la majorité des techniques d'imagerie actuelles l'amélioration ne peut passer que par des radiotélescopes de plus en plus 'grands'. Ceci est d'autant plus contraignant que les fréquences analysées sont basses. C'est une course au gigantisme à laquelle se livre la communauté des radioastronomes. De plus, ces techniques interférométriques nécessitent une déconvolution, étape extrêmement délicate, très sensible aux a priori et dont les performances impactent directement le *contraste* de l'image. Cette étape peut s'avérer irréversible dans la perte d'information sur l'image.

Spécifiquement, les performances des cartographies d'imagerie dynamique actuelles restent très largement en deçà de la sensibilité que permet en théorie la surface collectrice d'un instrument interférométrique tel que le SKA. En effet ses antennes très largement espacées rassemblent des milliers de capteurs et permettent des observations multi-fréquences, variant dans le temps. Néanmoins, les méthodes actuelles ne permettent pas d'exploiter ces diversités (espace, temps, fréquence), car cela soulève un problème méthodologique lié à la complexité/faisabilité du calcul en raison de la grande taille des données. De tels problèmes peuvent être par exemple traités par filtrages particuliers mais ceux-ci nécessitent de trop grandes capacités numériques. Par ailleurs, les méthodes basées sur les filtres de Kalman non linéaires ne sont pas assez précises pour prendre en compte de fortes non linéarités et voient de ce fait leurs performances se dégrader. Ainsi, la plupart des méthodes actuellement utilisées sont conçues en supposant que les paramètres d'intérêt (par exemple, la cohérence de la source) et/ou les paramètres auxiliaires (par exemple, ceux liés au déphasage) sont statiques au cours de l'observation [5-8]. Par conséquent, les données sont moyennées de manière approximatives sur des intervalles trop longs, ce qui conduit à une synthèse sous-optimale. De plus, la plupart de l'état de l'art suppose qu'il est possible de faire tenir l'ensemble des données traitées en mémoire et d'agréger toutes les données brutes dans la mémoire d'une seule unité de traitement, ce qui ne sera pas possible en pratique.

Il devient nécessaire de repousser les limitations des techniques actuelles en développant de nouvelles méthodes à la hauteur du formidable potentiel offert par SKA.

## Méthodologie et solutions proposées

Il nous semble dès lors pertinent de développer de nouvelles techniques d'imagerie dynamique basée sur le maximum de vraisemblance (MV). En effet le pouvoir de résolution de ces dernières n'est pas limité par l'étendue du réseau. De plus on peut aussi en améliorer le pouvoir de résolution en prenant en compte le flux des observations disponibles et en exploitant un modèle qui décrit la dynamique de l'image. Afin d'améliorer la résolution, le contraste des images et faire de l'imagerie dynamique, nous allons travailler sur des critères du maximum de vraisemblance paramétré ou non selon les *a priori* disponibles sur l'image à former [9]. Il semble évident que l'application brute-force du MV ne sera parfois ni envisageable à cause de données incomplètes ou encore ni compatible des puissances de calcul disponibles. Il est alors nécessaire de développer pour le maximum de vraisemblance des techniques de mises en œuvre dynamiques, éventuellement approchées, qui soient statistiquement performantes, temps réel, robustes et efficaces numériquement. Une alternative qui nous paraît prometteuse à étudier est l'algorithme EM (Expectation-maximization) dans sa version stochastique. En effet les algorithmes EM permettent d'accélérer la convergence quand on ne peut pas traiter le Maximum de vraisemblance directement. Afin de prendre en compte l'effet dynamique, nous adapterons le schéma susmentionné en un processus en ligne qui consiste à utiliser une combinaison d'une approximation stochastique de la vraisemblance avec un algorithme EM, visant à incorporer l'information apportée par la nouvelle observation disponible dans l'étape E et l'étape M. En particulier, l'utilisation de ce que l'on appelle l'approximation stochastique de l'EM (connu sous le nom de SAEM) semble être la meilleure alternative puisque son efficacité numérique a été démontrée [10].

Cette dernière sera soigneusement conçue en tempérant la SAEM afin d'échapper aux maxima locaux [11].

De plus en raison de la géométrie des stations, nous proposons d'effectuer l'étape E et l'étape M de la SAEM d'une manière distribuée, dans laquelle chaque station peut calculer indépendamment une forme compacte de quantités locales sommaires en utilisant uniquement des observations locales [12, 13, 14]. Ensuite le processus central estime les paramètres communs (c'est-à-dire les images) qui seront renvoyées à chaque station afin d'affiner le processus d'apprentissage. Il convient de mentionner que les quantités globales résumées seront considérées comme des moyennes pondérées des quantités locales résumées. Afin de gagner en robustesse, ces pondérations seront estimées localement en tenant compte de l'effet aléatoire local. Les algorithmes proposés seront systématiquement comparés aux méthodes de l'état de l'art et les performances statistiques obtenues seront comparées aux bornes minimales d'estimation [15].

L'efficacité opérationnelle des algorithmes précédents repose cependant sur la validité et la précision du modèle exploité dans les algorithmes. Nous proposerons pour ces divers algorithmes une mise en œuvre sous forme d'*algorithm unrolling* [16]. En effet l'utilisation de *model based deep learning techniques* assurera d'une part une vitesse de convergence supérieure et permettra de donner des degrés de liberté au modèle de réception postulé afin de tenir compte des problèmes de calibration. Le déroulement de l'algorithme itératif sur les couches successives du réseau de neurones bénéficiera de l'adjonction structurelle de paramètres supplémentaires liés aux connexions entre les neurones. Ces degrés de liberté supplémentaires pourront par exemple pallier la méconnaissance des paramètres de calibration (par exemple perturbations non maîtrisées liées à la traversée de l'ionosphère [17]). Il s'agit ici d'une alternative au *deep neural network* dont l'opacité et la non interprétabilité de l'apprentissage statistique peut s'avérer parfois déroutante. L'apprentissage évoluerait alors autour d'une boussole pointée et contrôlée par le modèle de l'observation. L'adaptation des paramètres supplémentaires serait seul sujet à un apprentissage statistique 'aveugle'. Il sera alors intéressant de procéder à une comparaison des méthodes de traitement de problème inverse entièrement fondés sur le seul modèle de réception avec celles fondées sur le *model based deep learning* et aussi celles sur le *total deep neural network*. Une telle comparaison apportera sa pierre à l'édifice dans le débat qui a lieu à ce sujet entre les communautés traitement du signal d'une part et apprentissage statistique de l'autre.

## Plan de travail

- 1 Bibliographie en traitement d'antenne, radioastronomie et modèle d'image.
- 2 Appropriation et étude de méthodes du maximum de vraisemblance pour une image statique
- 3 Proposition de modèles adaptés d'évolution de l'image compatible d'une approximation stochastique de la vraisemblance.
- 4 Développement de méthodes SAEM pour une imagerie dynamique
- 5 Proposition de mises en œuvre distribuées
- 6 Implémentation des méthodes dans un cadre de *model based deep learning*
- 7 Validation/Immersion dans un observatoire avec des radioastronomes

**Conditions de travail :** Le doctorant travaillera au sein laboratoire SATIE à l'ENS Paris-Saclay, 4 rue des sciences, 91190 Gif sur Yvette. Il intégrera l'équipe MOSS (Méthodes et Outils pour les Signaux et Systèmes). Il disposera d'un bureau au sein de l'activité traitement du signal pour la radioastronomie. Il participera à une école d'été sur la radioastronomie.

## Références

- [1] C. Ferrari, Le Square Kilometre Array (SKA) : un radiotélescope géant pour étudier l'aube et l'évolution du cosmos *Reflets phys.* N°67 (2020) 10-14
- [2] C. Ferrari et al., French SKA White Book - "The French community towards the Square Kilometre" Array ", Published by the SKA-France Coordination in collaboration with AS SKA-LOFAR, arXiv:1712.06950, 2017.
- [3] E. Boyer, P. Forster et P. Larzabal, "Non Asymptotic Performance Analysis of Beamforming with Stochastic Signals", IEEE Signal Processing letters. Volume 11, Number 1, pp. 23-25, Jan 2004.
- [4] S. J. Wijnholds and A.-J. van der Veen, "Fundamental imaging limits of radio telescope arrays" IEEE Journal of Selected Topics in Signal Processing, vol. 2, no. 5, pp. 613-623, 2008
- [5] J. Birdi, A. Repetti and Y Wiaux, "Polca SARA—full polarization, direction-dependent calibration, and sparse imaging for radio interferometry", Monthly Notices of the Royal Astronomical Society, pp. 3509-3528, March 2020
- [6] L Wei and S. J. Wijnholds, "Joint calibration and imaging for phased array radio telescopes," Monthly Notices of the Royal Astronomical Society, V483, no. 4, pp5672-5686, March 2019.
- [7] Wiaux, Y., Jacques, L., Puy, G., Scaife, A.M.M., Vanderghyest, P.: Compressed sensing imaging techniques for radio interferometry. Monthly Notices of the Royal Astronomical Society **395**, 1733–1742 (2009)
- [8] A.-J. van der Veen, S. J. Wijnholds and A. Sardarabadi "Signal processing for radio astronomy" in Handbook of Signal Processing Systems. Springer, 2013, pp. 421-463. 2003.
- [9] M Simeoni, P Hurley "SiML: Sieved Maximum Likelihood for Array Signal Processing" - arXiv preprint arXiv:2102.01950, 2021
- [10] M. Lavielle. Mixed Effects Models for the Population Approach: Models, Tasks, Methods and Tools. Chapman and Hall/CRC, 2014
- [11] S. Allasonnière and J. Chevallier, "A new class of stochastic EM algorithms. Escaping local maxima and handling intractable sampling", Computational Statistics & Data Analysis 159, 2021
- [12] A. Dieuleveut, G. Fort, E. Moulines, G. Robin. Federated Expectation Maximization with heterogeneity mitigation and variance reduction, NeurIPS, May 2021.
- [13] A. Hippert-Ferrer, M. N. El Korso, A. Breloy, G. Ginolhac, "Robust Mean and Covariance Matrix Estimation Under Heterogeneous Mixed-Effects Model with Missing Values", Signal Processing, Vol. 188, Nov. 2021, 108-195.

- [14] V. Ollier, M. N. El Korso, A. Ferrari, R. Boyer and P. Larzabal, "Robust distributed calibration of radio interferometers with direction dependent distortions", Elsevier Signal Processing Journal, Volume 153, December 2018, Pages 348-354
- [15] C Ren, J Galy, E Chaumette, F Vincent, P Larzabal, A Renaux, "Recursive Hybrid Cramér–Rao Bound for Discrete-Time Markovian Dynamic Systems "IEEE Signal Processing Letters 22 (10), Oct 2015, pp. 1543-1547
- [16] V. Monga, Y. Li, and Y. C. Eldar "Unrolling *Interpretable, efficient deep learning for signal and image processing*" IEEE signal processing magazine, March 2021, pp.18-44.
- [17] M. Brossard, M. N. El Korso, M. Pesavento, R. Boyer, P. Larzabal and S. Wijnholds, "Parallel Multi-Wavelength Calibration Algorithm for Radio Astronomical Arrays", Elsevier Signal Processing Journal, Volume 145, April 2018, pp. 258–271