

**Université de Montréal**

**Mesurer la masse de trous noirs supermassifs à l'aide  
de l'apprentissage automatique**

par

**David Chemaly**

Département de Physique  
Faculté des arts et des sciences

Mémoire présenté en vue de l'obtention du grade de  
Maître ès sciences (M.Sc.)  
en Astrophysique et Astronomie

26 juillet 2023

# Université de Montréal

Faculté des arts et des sciences

---

Ce mémoire intitulé

## Mesurer la masse de trous noirs supermassifs à l'aide de l'apprentissage automatique

présenté par

**David Chemaly**

a été évalué par un jury composé des personnes suivantes :

*Laurence Perreault-Levasseur*

---

(président-rapporteur)

*Yashar Hezaveh*

---

(directeur de recherche)

*Julie Hlavacek-Larrondo*

---

(codirecteur)

*Patrick Dufour*

---

(membre du jury)

# Résumé

---

Des percées récentes ont été faites dans l'étude des trous noirs supermassifs (SMBH), grâce en grande partie à l'équipe du télescope de l'horizon des événements (EHT). Cependant, déterminer la masse de ces entités colossales à des décalages vers le rouge élevés reste un défi de taille pour les astronomes. Il existe diverses méthodes directes et indirectes pour mesurer la masse de SMBHs. La méthode directe la plus précise consiste à résoudre la cinématique du gaz moléculaire, un traceur froid, dans la sphère d'influence (SOI) du SMBH. La SOI est définie comme la région où le potentiel gravitationnel du SMBH domine sur celui de la galaxie hôte. Par contre, puisque la masse d'un SMBH est négligeable face à la masse d'une galaxie, la SOI est, d'un point de vue astronomique, très petite, typiquement de quelques dizaines de parsecs. Par conséquent, il faut une très haute résolution spatiale pour étudier la SOI d'un SMBH et pouvoir adéquatement mesurer sa masse. C'est cette nécessité d'une haute résolution spatiale qui limite la mesure de masse de SMBHs à de plus grandes distances. Pour briser cette barrière, il nous faut donc trouver une manière d'améliorer la résolution spatiale d'objets observés à un plus au décalage vers le rouge.

Le phénomène des lentilles gravitationnelles fortes survient lorsqu'une source lumineuse en arrière-plan se trouve alignée avec un objet massif en avant-plan, le long de la ligne de visée d'un observateur. Cette disposition a pour conséquence de distordre l'image observée de la source en arrière-plan. Puisque cette distorsion est inconnue et non-linéaire, l'analyse de la source devient nettement plus complexe. Cependant, ce phénomène a également pour effet d'étirer, d'agrandir et d'amplifier l'image de la source, permettant ainsi de reconstituer la source avec une résolution spatiale considérablement améliorée, compte tenu de sa distance initiale par rapport à l'observateur.

L'objectif de ce projet consiste à développer une chaîne de simulations visant à étudier la faisabilité de la mesure de la masse d'un trou noir supermassif (SMBH) par cinématique du gaz moléculaire à un décalage vers le rouge plus élevé, en utilisant l'apprentissage automatique pour tirer parti du grossissement généré par la distorsion d'une forte lentille gravitationnelle. Pour ce faire, nous générons de manière réaliste des observations du gaz moléculaire obtenues par le Grand Réseau d'Antennes Millimétrique/Submillimétrique de

l'Atacama (ALMA). Ces données sont produites à partir de la suite de simulations hydrodynamiques *Rétroaction dans des Environnements Réalistes (FIRE)*. Dans chaque simulation, l'effet cinématique du SMBH est intégré, en supposant le gaz moléculaire virialisé. Ensuite, le flux d'émission du gaz moléculaire est calculé en fonction de sa vitesse, température, densité, fraction de  $H_2$ , décalage vers le rouge et taille dans le ciel. Le cube ALMA est généré en tenant compte de la résolution spatiale et spectrale, qui dépendent du nombre d'antennes, de leur configuration et du temps d'exposition. Finalement, l'effet de la forte lentille gravitationnelle est introduit par la rétro-propagation du faisceau lumineux en fonction du profil de masse de l'ellipsoïde isotherme singulière (SIE).

L'exploitation de ces données ALMA simulées est testée dans le cadre d'un problème de régression directe. Nous entraînons un réseau de neurones à convolution (CNN) à apprendre à prédire la masse d'un SMBH à partir des données simulées, sans prendre en compte l'effet de la lentille. Le réseau prédit la masse du SMBH ainsi que son incertitude, en supposant une distribution a posteriori gaussienne. Les résultats sont convaincants : plus la masse du SMBH est grande, plus la prédiction du réseau est précise et exacte. Tout comme avec les méthodes conventionnelles, le réseau est uniquement capable de prédire la masse du SMBH tant que la résolution spatiale des données permet de résoudre la SOI. De plus, les cartes de saillance du réseau confirment que celui-ci utilise l'information contenue dans la SOI pour prédire la masse du SMBH. Dans les travaux à venir, l'effet des lentilles gravitationnelles fortes sera introduit dans les données pour évaluer s'il devient possible de mesurer la masse de ces mêmes SMBHs, mais à un décalage vers le rouge plus élevé.

**Mots-clés : trous noirs supermassifs – apprentissage automatique – lentilles gravitationnelles – simulations hydrodynamiques – dynamiques du gaz moléculaire – sphère d'influence – interférométrie – réseau de neurone convolutif**



# Abstract

---

Recent breakthroughs have been made in the study of supermassive black holes (SMBHs), thanks largely to the Event Horizon Telescope (EHT) team. However, determining the mass of these colossal entities at high redshifts remains a major challenge for astronomers. There are various direct and indirect methods for measuring the mass of SMBHs. The most accurate direct method involves resolving the kinematics of the molecular gas, a cold tracer, in the SMBH's sphere of influence (SOI). The SOI is defined as the region where the gravitational potential of the SMBH dominates that of the host galaxy. However, since the mass of a SMBH is negligible compared to the mass of a galaxy, the SOI is, from an astronomical point of view, very small, typically a few tens of parsecs. As a result, very high spatial resolution is required to study the SOI of a SMBH and adequately measure its mass. It is this need for high spatial resolution that limits mass measurements of SMBHs at larger distances. To break this barrier, we need to find a way to improve the spatial resolution of objects observed at higher redshifts.

The phenomenon of strong gravitational lensing occurs when a light source in the background is aligned with a massive object in the foreground, along an observer's line of sight. This arrangement distorts the observed image of the background source. Since this distortion is unknown and non-linear, analysis of the source becomes considerably more complex. However, this phenomenon also has the effect of stretching, enlarging and amplifying the image of the source, enabling the source to be reconstructed with considerably improved spatial resolution, given its initial distance from the observer.

The aim of this project is to develop a chain of simulations to study the feasibility of measuring the mass of a supermassive black hole (SMBH) by kinematics of molecular gas at higher redshift, using machine learning to take advantage of the magnification generated by the distortion of a strong gravitational lens. To this end, we realistically generate observations of molecular gas obtained by the Atacama Large Millimeter/Submillimeter Antenna Array (ALMA). These data are generated from the Feedback in Realistic Environments (FIRE) suite of hydrodynamic simulations. In each simulation, the kinematic effect of the SMBH is integrated, assuming virialized molecular gas. Next, the emission flux of the molecular gas is calculated as a function of its velocity, temperature, density,  $\text{H}_2$  fraction, redshift and

sky size. The ALMA cube is generated taking into account spatial and spectral resolution, which depend on the number of antennas, their configuration and exposure time. Finally, the effect of strong gravitational lensing is introduced by back-propagating the light beam according to the mass profile of the singular isothermal ellipsoid (SIE).

The exploitation of these simulated ALMA data is tested in a direct regression problem. We train a convolution neural network (CNN) to learn to predict the mass of an SMBH from the simulated data, without taking into account the effect of the lens. The network predicts the mass of the SMBH as well as its uncertainty, assuming a Gaussian a posteriori distribution. The results are convincing: the greater the mass of the SMBH, the more precise and accurate the network's prediction. As with conventional methods, the network is only able to predict the mass of the SMBH as long as the spatial resolution of the data allows the SOI to be resolved. Furthermore, the network's saliency maps confirm that it uses the information contained in the SOI to predict the mass of the SMBH. In future work, the effect of strong gravitational lensing will be introduced into the data to assess whether it becomes possible to measure the mass of these same SMBHs, but at a higher redshift.

**Keywords:** supermassive black holes – machine learning – gravitational lensing – hydrodynamical simulations – molecular gas dynamics – sphere of influence – interferometry – convolutional neural network

# Table des matières

---

<b>Résumé</b> .....	v
<b>Abstract</b> .....	vii
<b>Table des figures</b> .....	xi
<b>Liste des sigles et des abréviations</b> .....	xiii
<b>Remerciements</b> .....	xvii
<b>Chapitre 1. Mise en contexte théorique</b> .....	1
1.1. Les galaxies .....	1
1.2. Les trous noirs .....	3
1.3. Liens entre les trous noirs supermassifs et leur galaxie hôte .....	6
1.4. Méthodes de mesure des masses de trous noirs supermassifs .....	8
1.4.1. Méthode directe : cinématique du gaz ionisé .....	8
1.4.2. Méthode directe : cinématique stellaire .....	10
1.4.3. Méthode directe : cinétique du disque mégamaser .....	10
1.4.4. Méthode directe : cinématique du gaz moléculaire .....	11
1.4.5. Méthode indirecte : plan fondamental de l'activité des trous noirs .....	13
1.4.6. Méthode indirecte : profil de brillance de surface .....	14
1.4.7. Méthode indirecte : cartographie par réverbération .....	14
1.5. Les simulations cosmologiques .....	16
1.5.1. Les simulations à N corps .....	17
1.5.2. Les simulations hydrodynamiques .....	20
1.6. Les lentilles gravitationnelles .....	22
1.6.1. L'équation de lentille .....	23
1.6.2. Potentielle de la lentille .....	27
1.6.3. Grossissement et étirement local .....	28

1.7.	L'apprentissage automatique .....	30
1.7.1.	Les réseaux de neurones artificiels .....	30
1.7.2.	Les réseaux de neurones convolutifs .....	34
1.7.3.	L'augmentation des données .....	35
1.7.4.	Les incertitudes .....	35
1.8.	Description de ce mémoire .....	36
1.8.1.	Objectifs .....	36
1.9.	Déclaration de l'étudiant .....	37
<b>Chapitre 2. Measuring the mass of supermassive black holes at higher redshifts through gravitational lensing with machine learning.</b>		<b>39</b>
	Abstract .....	40
2.1.	Introduction .....	40
2.2.	Data .....	43
2.2.1.	FIRE simulations .....	43
2.3.	Simulations .....	44
2.3.1.	Black Hole injection .....	44
2.3.2.	CO emission line .....	46
2.3.3.	Gravitationnal Lensing .....	48
2.3.4.	Mock ALMA data .....	50
2.3.5.	Data augmentation .....	53
2.4.	Neural Network .....	54
2.4.1.	Architecture .....	54
2.4.2.	Estimating Gaussian Uncertainties .....	56
2.5.	Discussion .....	57
2.6.	Conclusion .....	60
2.7.	Acknowledgements .....	61
<b>Chapitre 3. Conclusions</b> .....		<b>63</b>
	Bibliographie .....	66

# Table des figures

---

1.1	Comparaison entre une galaxie spirale et elliptique.....	2
1.2	Effet de la rotation d'un trou noir sur son disque d'accrétion.....	4
1.3	Imagerie directe des trous noirs M87 et Sagittarius A*.....	5
1.4	Relation entre la masse de SMBHs et leur galaxie hôte.....	7
1.5	Vitesse du gaz moléculaire en fonction de la position radiale du trou noir supermassif.....	12
1.6	Plan fondamental de l'activité des trous noirs.....	15
1.7	Exemple d'une galaxie générée par la simulation FIRE.....	22
1.8	Image d'une lentille gravitationnelle prise par Hubble.....	23
1.9	Schéma représentatif de l'effet d'une lentille gravitationnelle.....	24
1.10	Schéma représentatif d'un réseau de neurones artificiels.....	32
1.11	Schéma représentatif d'un réseau de neurones convolutifs.....	35
2.1	Density profil of a simulated galaxy from MassiveFIRE.....	44
2.2	PVD diagram of the same galaxy with varying SMBH masses.....	46
2.3	Impact of a masked Fourier transform on a clean image.....	52
2.4	Representation of a simulated ALMA cube.....	53
2.5	Schematic representation of the CNN.....	55
2.6	Network's predictions of the SMBH mass compared to the true values.....	58
2.7	Unbiased predictions of heavy SMBH masses.....	59
2.8	Examples of saliency maps from various galaxies.....	60

## Liste des sigles et des abréviations

---

AI	Intelligence artificielle, de l'anglais <i>Artificial Intelligence</i>
ALMA	Grand réseau d'antennes millimétrique/submillimétrique de l'Atacama, de l'anglais <i>Atacama Large Millimeter/submillimeter Array</i>
ANN	Réseau de neurone artificiel, de l'anglais <i>Artificial Neural Network</i>
CFHT	Télescope Canada France Hawaï, de l'anglais <i>Canada France Hawaii Telescope</i>
CNN	Réseau de neurone convolutif, de l'anglais <i>Convolutional Neural Network</i>
EHT	Télescope de l'horizon des évènements, de l'anglais <i>Event Horizon Telescope</i>
ESA	Agence spatiale européenne, de l'anglais <i>European Space Agency</i>
FWHM	Largeur à mi-hauteur, de l'anglais <i>Full Width Half Maximum</i>

FIRE	Rétroaction dans des environnements réalistes, de l'anglais <i>Feedback In Realistic Environment</i>
HST	Télescope spatiale Hubble, de l'anglais <i>Hubble Space Telescope</i>
ISCO	Orbite interne la plus stable, de l'anglais <i>Inner Most Stable Circular Orbit</i>
ISM	Milieu Interstellaire, de l'anglais <i>interstellar medium</i> )
JPL-Caltech	Laboratoire de jet à propulsion de institut de technologie de Californie, de l'anglais <i>Jet Propulsion Laboratory at CalTech</i>
kpc	Kiloparsec
LIGO	Observatoire d'ondes gravitationnelles par interférométrie laser, de l'anglais <i>Laser Interferometer Gravitational-Wave Observatory</i>
LSST	Grand télescope d'étude synoptique, de l'anglais Large Synoptic Survey Telescope
MCMC	Méthode de Monte-Carlo par chaînes de Markov, de l'anglais <i>Markov Chain Monte Carlo</i>
MNRAS	Comptes rendus mensuels de la Société royale d'astronomie, de l'anglais <i>Monthly Notices of the Royal Astronomical Society</i>

Mpc	Mégaparsec
MSE	Erreur quadratique moyenne, de l'anglais <i>Mean Squared Error</i>
NASA	Aéronautique nationale et administration de l'espace, de l'anglais <i>National Aeronautics and Space Administration</i>
FIRE	Rétroaction dans des environnements réalistes, de l'anglais <i>Feedback In Realistic Environments</i>
NFW	Navarro Frenk White
NIR	Infrarouge proche, de l'anglais <i>Near-Infrared</i>
PP	Particule à particule, de l'anglais <i>Particle-Particle</i>
PM	Particule à maillage, de l'anglais <i>Particle-Mesh</i>
P3M	Particule à particule à particule à maillage, de l'anglais <i>Particle-Particle-Particle-Mesh</i>
SGD	Descente du gradient stochastique, de l'anglais <i>Stochastic Gradient Descent</i>
SMBH	Trou noir supermassif, de l'anglais <i>Supermassive Black Hole</i>



SOI	Sphère d'influence, de l'anglais <i>Sphere of Influence</i>
SPH	Hydrodynamique des particules lissées, de l'anglais <i>Smoothed Particle Hydrodynamics</i>
STIS	Spectrographe d'imagerie du télescope spatial Hubble, de l'anglais <i>Space Telescope Imaging Spectrograph</i>
UV	Ultraviolet
VLBI	Interférométrie à très longue base, de l'anglais <i>Very Long Baseline Interferometry</i>
WISDOM	Relevé interférométrique à ondes millimétriques des masses d'objets sombres, de l'anglais <i>mm-Wave Interferometric Survey of Dark Object Masses</i>
$\lambda$ CDM	Matière noire froide avec constante cosmologique, de l'anglais <i>lambda cold dark matter</i>

# Remerciements

---

J'aimerais tout d'abord commencer par exprimer ma profonde gratitude envers mes superviseurs, Julie Hlavacek-Larrondo et Yashar Hezaveh. Votre soutien et mentorat constants tout au long de ma maîtrise m'ont permis de surmonter les nombreux obstacles qui se sont dressés sur mon chemin. Je tiens à vous remercier non seulement pour votre précieuse aide, mais aussi pour votre gentillesse et votre sens de l'humour qui ont rendu cette expérience encore plus enrichissante.

Je souhaite également exprimer ma reconnaissance envers tous les membres du groupe de recherche X-Tra, ainsi qu'à ceux de l'institut récemment créé, Ciela. Vos sourires et nos conversations passionnantes m'ont donné la motivation nécessaire pour me lever chaque matin et venir travailler au MIL, alors que j'aurais facilement pu poursuivre mes travaux à distance, dans le confort de ma chambre. Votre présence et votre dynamisme ont indéniablement ajouté une dimension humaine précieuse à mon parcours académique.

Enfin, je ne peux terminer ces remerciements sans mentionner ma famille et mes amis en dehors du milieu académique. Sans eux, j'aurais certainement succombé aux abîmes obsessionnels de la recherche. Leur soutien inconditionnel, leur amour et leurs encouragements ont été la force motrice qui m'a permis d'aller de l'avant, en dépit des défis et des moments de doute. Je leur suis profondément reconnaissant de m'avoir accompagné tout au long de cette aventure.

# Chapitre 1

---

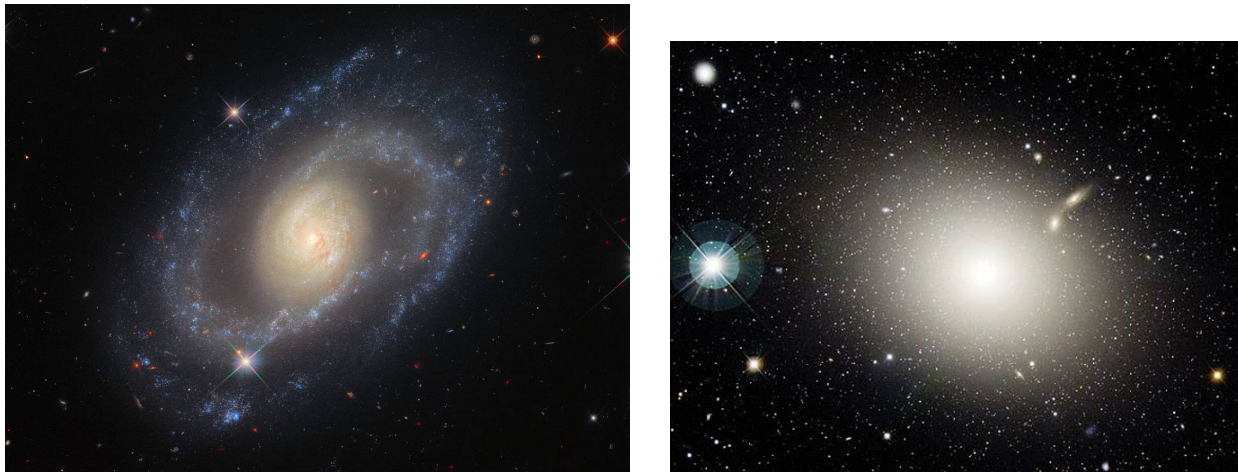
## Mise en contexte théorique

### 1.1. Les galaxies

L'Univers est composé de deux types de matière : la matière baryonique, constituée d'électrons, de protons et d'autres particules, et la matière sombre, qui ne semble aucunement interagir avec la lumière, ni la matière baryonique. Les galaxies peuvent simplement être considérées comme des amas de gaz qui, au fil du temps, s'effondrent sous l'attraction gravitationnelle d'un halo de matière sombre. Ce processus de condensation conduit à la formation stellaire (Caldwell & Ostriker 1981).

Les galaxies sont principalement divisées en 2 grandes familles : les spirales et les elliptiques. Les galaxies spirales, originellement décrites par Edwin Hubble en 1936, consistent en un disque rotatif plat contenant du gaz, des étoiles et de la poussière avec un centre très dense nommée le bulbe (Hubble 1936). Les galaxies spirales prennent leur nom des bras spiraux qui se répandent du bulbe au bout du disque galactique (voir Figure 1.1 à gauche). Puisque les bras spiraux sont des zones de surdensité, le taux de formation stellaire au sein de ceux-ci est beaucoup plus élevé que dans le reste du halo stellaire (Mihalas & Routly 1968). Un haut taux de formation stellaire mène donc à une plus grande population d'étoiles jeunes expliquant par le fait même pourquoi les bras spiraux sont si lumineux. Ces régions à haut taux de formation stellaire donne aux galaxies spirales une teinte bleuâtre. Les galaxies spirales s'étendent sur une large gamme de tailles, de 5 à 100 kpc, de masse,  $10^9$  à  $10^{12} M_{\odot}$  et de luminosité,  $10^8$  à  $10^{11} L_{\odot}$ . Il a aussi été observé qu'environ deux tiers des galaxies spirales ont une composante supplémentaire sous la forme d'une structure en barre s'étendant du bulbe au début des bras spiral (Benjamin et al. 2005).

Quant à elles, les galaxies elliptiques sont moins complexes. Elles n'ont pas de bras, mais plutôt une émission lumineuse diffuse qui leur donne la forme d'une ellipse (d'où le nom). Une galaxie elliptique à un faible taux de formation stellaire et une population d'étoiles vieillissante qui peut généralement être décrite par une population stellaire d'âge unique (Dressler 1980). N'étant généralement pas en train de former d'étoiles, les galaxies elliptiques



**Figure 1.1** – Comparaison entre les 2 grandes familles de galaxies mettant de l’avant la diversité de ces objets. **Gauche** : Mrk1337 est un exemple de galaxie spirale. La structure des bras spiraux où la formation stellaire est élevée est visible. Crédit : ESA/Hubble NASA. **Droite** : M87, une galaxie elliptique, caractérisée par une distribution lumineuse diffuse et homogène donnant à la galaxie une forme elliptique. Crédit : Canada-France-Hawaii Telescope (CFHT).

paraissent rouges. Les galaxies elliptiques varient considérablement en diamètre, 1 à 120 kpc, en masse,  $10^5$  à  $10^{12} M_{\odot}$ , et en luminosité,  $10^6$  à  $10^{12} L_{\odot}$  (Loveday 1996). Le profil de brillance de surface d’une galaxie elliptique peut facilement être décrit par le profil de Sérsic 1963,

$$I(R) = I_0 \exp \left[ -b_n \left( \frac{r}{R_e} \right)^{\frac{1}{n}} \right], \quad (1.1.1)$$

où  $I_0$  est l’intensité de surface centrale,  $n$ , l’indice de Sérsic responsable de la concentration lumineuse, et  $R_e$ , le rayon effectif, soit le rayon qui comprend la moitié de la luminosité totale.

Bien qu’il existe un grand nombre de galaxies cataloguées, la communauté scientifique en connaît peu sur leur évolution détaillée au fil du temps cosmique. Cela est majoritairement dû au fait que l’échelle temporelle à laquelle évolue une galaxie est de plusieurs ordres de grandeur plus élevé que l’échelle humaine. En d’autres mots, bien qu’une galaxie évolue au court du temps, il nous est impossible d’observer de manière directe cette évolution. Par contre, puisque la lumière à une vitesse fixe, toutes les galaxies observées représentent une image à différentes époques de l’Univers en fonction de leur distance. Nous avons ainsi accès à l’information évolutive des galaxies en termes de population, mais l’évolution temporelle d’une galaxie en particulier reste encore grandement incertaine.

Bien que l'étude des différentes catégories de galaxies ainsi que leur évolution soit fascinante, ce mémoire portera principalement sur les trous noirs supermassifs qui se trouvent en leur centre.

## 1.2. Les trous noirs

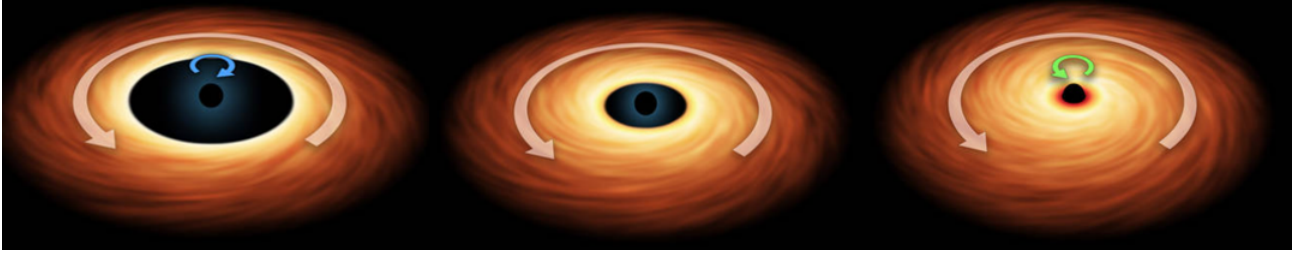
Selon la théorie de la relativité d'Einstein (Einstein 1905), l'attraction gravitationnelle d'une masse est simplement due à la courbure de l'espace-temps. Dans le cas d'un trou noir, le corps est si dense et compact que plus rien ne peut échapper à cette courbure - pas même la lumière. Les trous noirs prennent donc leur nom littéralement du fait qu'il s'agit d'une sphère sans lumière.

Bien que l'impact de ces objets sur leur environnement soit intense et complexe, un trou noir est simplement défini par 3 paramètres : sa charge, son moment angulaire ( $a$ ) et sa masse ( $M_{BH}$ ). Puisque l'Univers est considéré comme étant globalement neutre, la charge d'un trou noir se neutralise lors de sa formation (autant de particules chargées positivement ou négativement seront attirées). Ainsi, dans un contexte astrophysique, un trou noir est purement caractérisé par sa masse et son moment angulaire.

C'est Schwarzschild (Schwarzschild 1916) qui publie la métrique paramétrant les coordonnées de l'espace-temps affecté par un trou noir sans rotation. Cette métrique est aussi communément appelée «la métrique de Schwarzschild» :

$$ds^2 = - \left(1 - \frac{2GM_{BH}}{c^2 r}\right) c^2 dt^2 + \left(1 - \frac{2GM_{BH}}{c^2 r}\right)^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2, \quad (1.2.1)$$

ici  $G$  représente la constante gravitationnelle et  $ds$  peut simplement être considéré comme le quadrivecteur de position en coordonnées sphériques. Ainsi, chaque terme devant  $dt$ ,  $dr$ ,  $d\theta$  et  $d\phi$  représente respectivement les coordonnées temporelle, radiale et angulaires. De cette métrique, plusieurs phénomènes physiques fascinant à propos des trous noirs peuvent être déduit notamment l'apparition de deux autres rayons critiques à  $r = 0$  et  $r_s = \frac{2GM_{BH}}{c^2 r}$ . Ces valeurs sont critiques puisqu'elles mènent à une division par 0. À  $r = 0$ , soit le centre du trou noir, cette métrique prédit une courbure infinie en un point ponctuel : la singularité. En ce qui concerne le rayon Schwarzschild ( $r_s$ ), celui-ci correspond physiquement à la limite d'orbite stable autour du trou noir sans rotation. Toute orbite avec  $r < r_s$  sera donc instable et perdra en distance radiale avec le trou noir au fil du temps. Ce rayon est couramment nommé ISCO soit l'orbite interne la plus stable (de l'anglais *inner most stable circular orbit*).



**Figure 1.2** – Schématisation du rayon minimal d’orbite stable d’un trou noir en fonction de la rotation de celui-ci et du disque d’accrétion. Plus le trou noir à un moment angulaire élevé et plus l’horizon diminue si tout tourne dans la même direction (droite), inversement si les rotations sont opposées (gauche). Le trou noir du milieu n’a pas de rotation et fait simplement office de repère pour la comparaison. Crédit : NASA/JPL-Caltech.

Le rayon de Schwarzschild dépend uniquement de la masse du trou noir. Une fois ce rayon traversé, il n’y a plus de retour. Par contre, si le trou est en rotation alors la métrique de Schwarzschild n’est plus valide (voir Figure 1.2). Effectivement, si la matière orbitant le trou noir tourne dans la même direction que celui-ci alors le rayon minimal d’orbite stable sera inversement proportionnel au moment angulaire du trou noir. Au contraire, si la rotation de la matière orbitant le trou noir est inverse à celui-ci, la relation entre moment angulaire et ISCO sera directement proportionnel. Dans le cas d’un trou noir en rotation, la courbure de l’espace-temps environnante est paramétrée par la métrique de Kerr 1963 :

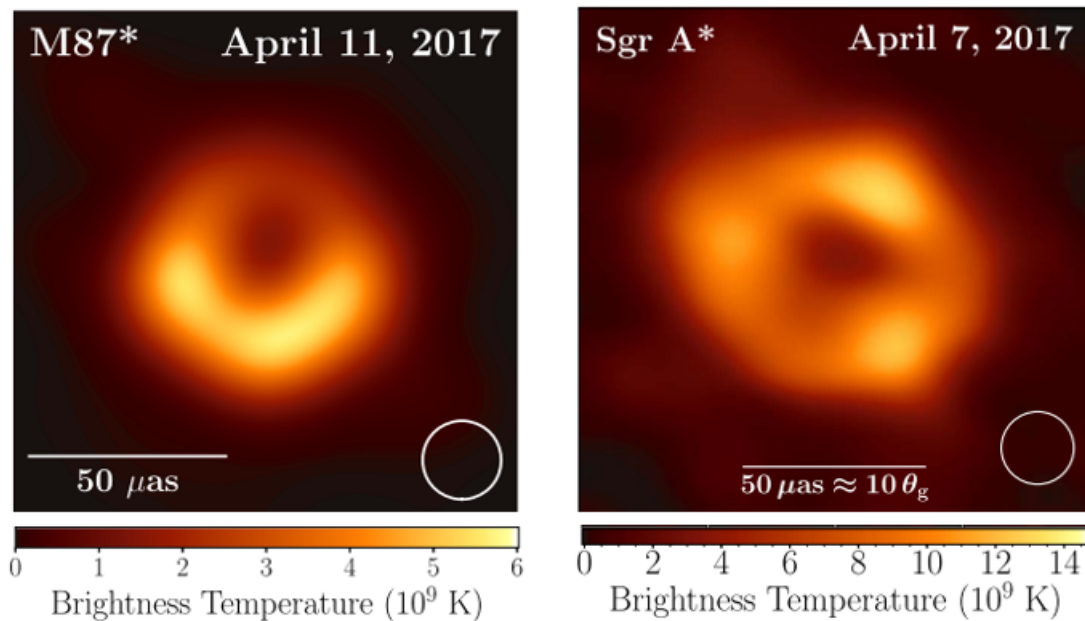
$$ds^2 = - \left( 1 - \frac{r_s}{r} + \frac{a^2}{r^2} \sin^2 \theta \right) c^2 dt^2 - \frac{r_s a}{r} \sin^2 \theta dt d\phi + \left( 1 - \frac{r_s}{r} + \frac{a^2}{r^2} \right)^{-1} dr^2 + r^2 d\theta^2 + \left( r^2 + \frac{r_s a^2}{r} \sin^2 \theta \right) \sin^2 \theta d\phi^2, \quad (1.2.2)$$

où  $a = \frac{J}{M_{BH}c}$  encode l’information du moment angulaire  $J$ . Pour un moment angulaire nul,  $J = 0$ , la métrique de Kerr revient à celle de Schwarzschild.

Les trous noirs sont souvent divisés en 3 catégories définis par leur masse : stellaire, intermédiaire et supermassif. Les trous noirs stellaires sont ceux avec la masse mesurée la plus faible ( $\lesssim 100 M_\odot$ ). Les trous noirs intermédiaires (IMBH – de l’anglais *Intermediate Mass Black Hole*) eux vont de  $\sim 100$  à  $10^6 M_\odot$ . Pour finir, les trous noirs supermassifs ont une masse de  $> 10^6 M_\odot$  (SMBH – de l’anglais *Supermassive Black Hole*) et seront l’objet principal de ce mémoire, se trouvent au centre de la vaste majorité des galaxies. Il est intéressant d’étudier ces SMBHs puisque plusieurs mesures semblent indiquer une corrélation entre la masse de ces trous noirs et diverses propriétés à grande échelle de leur galaxie hôte.

En avril 2019, le Télescope de l'Horizon des Événements (EHT – de l'anglais *Event Horizon Telescope*) a réalisé une percée historique en obtenant, pour la première fois, une image de l'horizon des événements d'un trou noir ainsi que du gaz environnant qui forme le disque d'accrétion (Figure 1.3). L'objet de cette observation n'était autre que le SMBH niché au cœur de la galaxie M87, localisée dans l'amas de galaxies de la Vierge (Event Horizon Telescope Collaboration et al. 2019).

Trois années plus tard, en avril 2022, l'EHT a publié une image de Sagittarius A\*, le SMBH situé au centre de notre propre galaxie, la Voie Lactée (Markoff & Event Horizon Telescope Collaboration 2022). Il est à noter que malgré la proximité relative de Sagittarius A\* par rapport à la Terre, sa masse inférieure rend sa taille angulaire similaire à celle de M87. De surcroît, sa masse réduite induit une vitesse orbitale du disque d'accrétion de quelques minutes seulement, comparée à la vitesse orbitale de plusieurs jours pour M87. Compte tenu de la complexité inhérente au traitement de données à caractère aussi variable, l'EHT a d'abord choisi de traiter les données de M87 puis celles de Sagittarius A\*, bien que celui-ci ait été observé en premier.



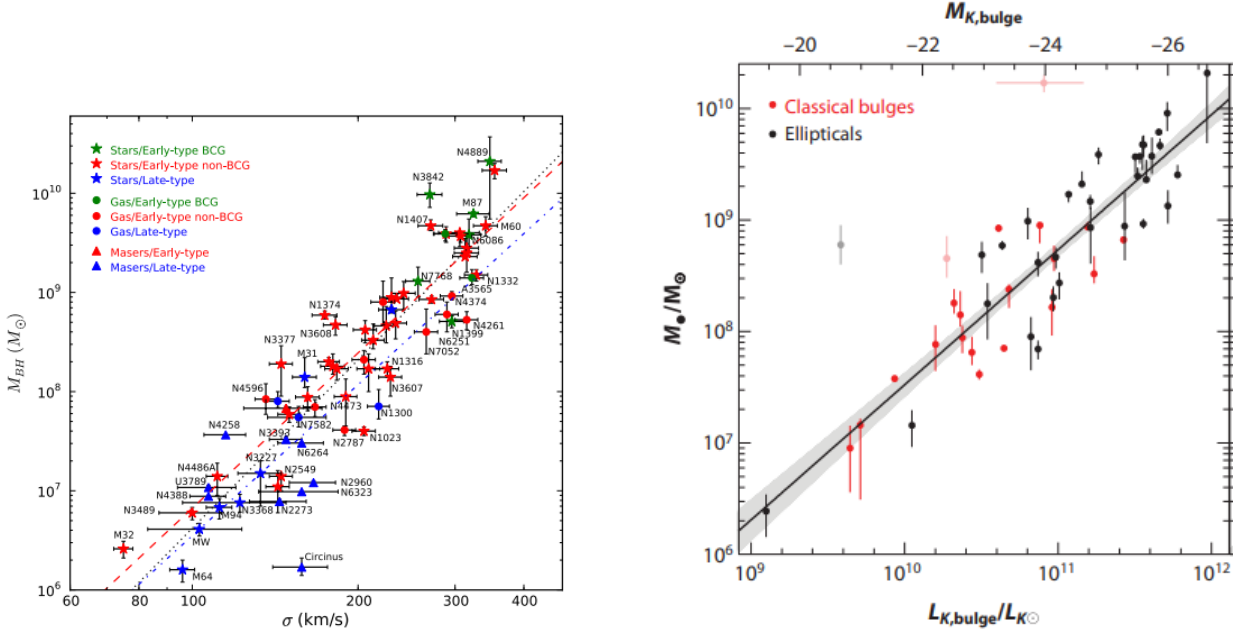
**Figure 1.3** – Les deux premières et seules observations directes de l'horizon d'un trou noir supermassif et son disque d'accrétion pris par le EHT. **Gauche** : Le trou noir au centre de M87. L'anneau de lumière représente le disque de matière orbitant autour du trou noir. Crédit : Event Horizon Telescope Collaboration et al. 2019. **Droite** : Image de Sagittarius A\*, le trou noir au centre de la Voie lactée. Crédit : Markoff & Event Horizon Telescope Collaboration 2022.

### 1.3. Liens entre les trous noirs supermassifs et leur galaxie hôte

Plusieurs observations de SMBHs au centre de galaxies suggèrent que leur masse ( $M_{\text{BH}}$ ) corrèle grandement avec plusieurs propriétés à grande échelle de leur galaxie hôte, notamment avec la dispersion de vitesse des étoiles,  $\sigma_*$ , mesurée à l'intérieur du rayon effectif englobant la moitié de la lumière totale émise par le bulbe galactique (e.g. Ferrarese & Merritt 2000b; Gebhardt et al. 2000b; Tremaine et al. 2002; Gültekin 2009), ainsi qu'avec la luminosité des étoiles dans le bulbe de la galaxie dans la bande K,  $L_{K,\text{bulbe}}$  (e.g. Dressler 1989; Kormendy & Richstone 1995; Marconi 2003; Sani et al. 2011). À noter que  $\sigma_*$  est souvent utilisé comme un analogue pour la masse de la galaxie puisque ces 2 variables sont directement proportionnelles. Ces relations observables impliquent une formation et une évolution commune entre les SMBHs et leur galaxie hôte (Silk 1998), ce qui est inattendu considérant le ratio des tailles (environ neuf ordres de grandeur de différence) et considérant que la contribution du SMBH au potentiel gravitationnel de la galaxie est négligeable. La nature de cette corrélation est non seulement incongrue, mais reste jusqu'à ce jour inexplicée. Ces deux relations locales,  $M_{\text{BH}} - \sigma_*$  et  $M_{\text{BH}} - L_{K,\text{bulbe}}$ , montrées à la Figure 1.4, sont utilisées pour prédire la masse des trous noirs dans les galaxies de tout type (McConnell & Ma 2013; Kormendy & Ho 2013b; Sexton et al. 2019). Par contre, ces deux relations font des prédictions très différentes (un ordre de magnitude à une masse  $\geq 10^{10} M_{\odot}$ , Lauer et al. 2007b) pour les trous noirs dans les galaxies les plus lumineuses et avec les plus grandes dispersions de vitesses (Lauer et al. 2007a; Bernardi et al. 2007; Shankar et al. 2016). Le nombre croissant de mesures de SMBHs a permis à la communauté scientifique de constater que les corrélations sont plus complexes que ce qui était initialement envisagé, ce qui peut être attribué, par exemple, à la diversité des types de galaxies (Kormendy & Ho 2013b). Aussi, le nombre de données à faible ( $\sim 10^6 M_{\odot}$ ) et à hautes ( $\gtrsim 10^9 M_{\odot}$ ) masses est très limité (voir Figure 1.4), ce qui réduit notre compréhension de la co-évolution des trous noirs et de leur galaxie hôte (e.g. Gültekin 2009; Graham & Scott 2013; McConnell & Ma 2013). Greene et al. 2010 et Läscher et al. 2016 ont trouvé que les SMBHs à faible masse se situent sous la relation standard  $M_{\text{BH}} - \sigma_*$ . De plus, en se basant sur des observations, il semblerait que les SMBHs avec une masse de plus de  $10^{10} M_{\odot}$  ne suivent pas ces relations puisqu'elles sous-estiment leur masse (voir Figure 1.4, e.g. McConnell et al. 2011b; Hlavacek-Larrondo et al. 2012b; Mezcua et al. 2018b; Mehrgan et al. 2019). Pour mieux comprendre ces relations, notamment leur pente et la dispersion intrinsèque des points, qui sont abondamment utilisées pour étudier l'évolution des galaxies, il faudrait un plus grand nombre de mesures précises de masses de SMBH surtout aux extrémités de la distribution.

Il existe plusieurs hypothèses cherchant à expliquer la physique derrière cette corrélation, mais rien n'a encore été établi (Heckman & Best 2014). Une idée serait que les SMBHs et





**Figure 1.4 – Gauche :** Relation entre la masse mesurée des SMBHs ( $M_{BH}$ ) et la dispersion de vitesse de la galaxie hôte à l’intérieur du rayon effectif englobant la moitié de la lumière totale émise par le bulbe galactique pour 72 galaxies. Les points verts sont les galaxies elliptiques dominantes des amas de galaxies, les points rouges sont les autres galaxies elliptiques et les points bleus sont les galaxies spirales. Il est indiqué aussi la méthode de mesure de la masse, soit par dynamique des étoiles (étoiles), du gaz (cercle) ou des mégamaser (triangles). Les meilleurs ajustements pour l’échantillon complet (noir), les galaxies elliptiques (rouge) et les galaxies spirales (bleu) sont aussi montrés. Crédit : McConnell & Ma 2013. **Droite :** Relation entre la masse mesurée des SMBHs ( $M_\bullet$ ) et la luminosité ( $L_{K,bulbe}$ ) et la magnitude absolue ( $M_{K,bulbe}$ ) des galaxies en bande K. Les galaxies elliptiques (en noir) et les galaxies avec des bulbes classiques, soit similaires aux galaxies elliptiques, mais avec une composante de disque (en rouge), sont indiquées. L’ajustement n’inclue pas les points pâles. Crédit : Kormendy & Ho 2013b.

leur galaxie hôte naissent tous avec plus ou moins la même masse et qu’ils gagnent en masse en se fusionnant entre eux au fil du temps (Schawinski et al. 2010). Une façon de valider ou réfuter cette théorie consisterait à vérifier si ces relations demeurent valables à de grandes distances, lorsque l’Univers est plus jeune. Malheureusement, en raison de la nécessité d’une résolution spatiale élevée, la mesure de la masse de SMBHs et de ces relations sont limitées à l’Univers local (McLure & Jarvis 2002; Pensabene et al. 2020). Dans la section suivante, les différentes méthodes de mesure des masses de trous noirs ainsi que leurs avantages et inconvénients seront discutés.

## 1.4. Méthodes de mesure des masses de trous noirs supermassifs

La masse de SMBHs peut soit être mesurée directement ou indirectement. Une mesure directe est beaucoup limitée puisqu'elle demande une haute résolution spatiale, mais permet une réelle mesure exacte et précise du SMBH en question. Dans le cas d'une mesure indirecte, les limitations sont moindres, ce qui facilite l'obtention approximative de la masse du SMBH. En contrepartie, la valeur obtenue est très incertaine et peut même être inexacte. Il existe plusieurs méthodes de mesure directe et indirecte de la masse des SMBHs. Nous examinerons certaines d'entre elles dans les sous-sections suivantes.

### 1.4.1. Méthode directe : cinématique du gaz ionisé

La détermination de la masse d'un SMBH implique la détection de sa signature cinématique au sein de sa galaxie hôte. Cette détection peut être réalisée en déterminant spatialement la cinématique à l'intérieur de la sphère d'influence (SOI – de l'anglais *Sphere of Influence*) du trou noir. La SOI est la région où le potentiel gravitationnel du trou noir domine sur celui de la galaxie hôte. Son rayon ( $r_g$ ) est donné par :

$$r_g \approx \frac{GM_{\text{BH}}}{\sigma_*^2}. \quad (1.4.1)$$

Par contre, même pour les galaxies elliptiques les plus lumineuses qui contiennent habituellement des SMBHs de masses  $> 10^8 M_\odot$ ,  $r_g$  est habituellement de l'ordre de quelques dizaines de parsecs soit uniquement  $\sim 0,1''$  à une distance de 20 Mpc. Cela est dû au fait que la masse d'un SMBH est négligeable face à celle d'une galaxie. Ainsi, étudier la cinématique d'étoiles ou de gaz à l'intérieur de  $r_g$  est très difficile à mesurer.

La majorité des mesures de masses de trous noirs ( $\sim 200$  mesures jusqu'à présent) est faite via la cinématique d'étoiles (voir Section 1.4.2, e.g. van den Bosch & de Zeeuw 2010; McConnell et al. 2012; Rusli et al. 2013) ou du gaz atomique ionisé (van der Marel & van den Bosch 1998; Barth et al. 2001) à l'aide d'observations spectroscopiques à l'intérieur de  $r_g$ . Étant donné que la cinématique à l'intérieur de  $r_g$  n'est pas complètement déterminée en raison des limites de résolution spatiale (seulement quelques mesures dans l'intervalle), les résultats obtenus à l'aide de ces techniques présentent des incertitudes statistiques qui peuvent entraîner une erreur dans la détermination de la masse, pouvant aller jusqu'à un

facteur 2 (van den Bosch & de Zeeuw 2010). En outre, lorsque les deux méthodes de modélisation dynamique sont applicables à une même galaxie, les valeurs de la masse obtenues ne sont souvent pas en accord et peuvent varier d’un facteur de 2 à 4 (Kormendy & Ho 2013b).

Pour mesurer la masse de SMBH en utilisant la cinématique du gaz ionisé dans la sphère d’influence, il faut avoir des observations spectroscopiques à différentes positions autour du trou noir. Par exemple, Walsh et al. 2013 ont obtenu des données du télescope spatial Hubble (HST – de l’anglais *Hubble Space Telescope*) pour mesurer la masse du SMBH au centre de la galaxie M87. Des observations avec le spectrographe d’imagerie (STIS – de l’anglais *Space Telescope Imaging Spectrograph*) du télescope spatial HST ont été utilisées. Pour mesurer la masse d’un SMBH avec STIS par la cinématique du gaz ionisé, il faut plusieurs observations en fente côte à côte pour avoir des spectres qui couvre toute la sphère d’influence. Une gaussienne est ensuite ajustée simultanément à chaque ligne d’émission assez brillante (e.g. la raie  $H_\alpha$  ou  $[N\ II]\lambda\lambda 6548,6583$ ). Les paramètres mathématiques de ces gaussiennes nous permettent d’en apprendre sur la physique du gaz ionisé. En corrigeant pour le décalage de Doppler dû à l’expansion de l’Univers régie par la constante de Hubble ( $H_0$ , Hubble 1929), la position de la raie nous donne la vitesse dans la ligne de visée. Cette correction n’est nécessaire que si l’objet étudié se situe à l’extérieur du groupe local. En effet, à l’intérieur de celui-ci, la vitesse intrinsèque des objets domine sur la vitesse dû à l’expansion de l’Univers. Quant à la largeur à mi-hauteur du maximum d’intensité (FWHM – de l’anglais *Full Width Half Maximum*) d’une raie et l’aire sous la gaussienne, celles-ci corrélerent respectivement avec la dispersion de la vitesse et le flux. En répétant le même processus pour tous les pixels des fentes observées, il est possible de générer un champ de vitesse autour du SMBH.

Théoriquement, la masse d’un SMBH peut être déduite directement à partir du champ de vitesse du gaz ionisé (Fabian & Lasenby 2015). Toutefois, l’application de cette méthode présente une complexité significative, car le champ de vitesse observé est affecté non seulement par le SMBH, mais également par divers autres facteurs, incluant le profil de masse des étoiles, le ratio masse-lumière, et l’angle d’inclinaison de la galaxie (Barth et al. 2001). Afin d’identifier la masse du SMBH, il est donc impératif de reproduire le plus fidèlement possible les observations, en l’occurrence le champ de vitesse, en se fondant sur un modèle physique qui intègre l’ensemble des paramètres pertinents. Pour ce faire, un large éventail de combinaisons de paramètres est testé jusqu’à l’obtention d’un résultat inférieur à un seuil d’erreur préétabli. Plusieurs techniques existent pour optimiser cet échantillonnage de l’espace des paramètres, l’une des plus populaires étant la méthode de Monte-Carlo par chaînes de Markov (MCMC – de l’anglais *Markov Chain Monte Carlo*). Le modèle qui coïncide le mieux avec les observations est par la suite retenu, permettant, par le fait même, la détermination de la masse du SMBH (Gow et al. 2022).

Bien que la méthode de modélisation de la cinématique du gaz ionisé soit conceptuellement simple et directe, plusieurs problèmes majeurs peuvent considérablement compliquer

l’analyse. D’une part, le gaz ionisé est souvent soumis à un mouvement non circulaire, rendant l’équation de conversion de la vitesse circulaire à la masse du SMBH imparfaite (Fabian & Lasenby 2015). D’autre part, il existe d’importantes turbulences dans la dispersion de vitesse, nécessitant l’intégration de la pression qui soutient le disque de gaz dans la modélisation de la cinématique, ce qui exige des approximations (Barth et al. 2001). Il est également fort probable que de la poussière soit présente dans le gaz, ce qui pourrait masquer une partie du gaz ionisé central et modifier la mesure de la masse du trou noir en affectant la contribution stellaire incluse dans les modèles (Gow et al. 2022).

### 1.4.2. Méthode directe : cinématique stellaire

La méthode de modélisation de la cinématique stellaire s’apparente à celle du gaz ionisé (Section 1.4.1), avec cependant quelques distinctions notables. Plutôt que de se baser sur les lignes d’émission de  $H\alpha\lambda 6548 \text{ \AA}$  et de  $[N,II]\lambda\lambda 6548,6583 \text{ \AA}$ , cette méthode utilise les caractéristiques d’absorption stellaires issues du  $(2 - 0)^{12}\text{CO}$  et du  $(3 - 1)^{12}\text{CO}$ . De plus, durant la modélisation, il est essentiel de prendre en compte le potentiel gravitationnel de la galaxie, résultant de l’interaction entre le SMBH, les étoiles, et un halo de matière sombre de Navarro-Frenk-White (NFW, Navarro et al. 1996). Cependant, cette modélisation des orbites stellaires à l’échelle de la galaxie peut être confrontée à diverses incertitudes liées notamment au halo de matière sombre (Gultekin 2009), à la structure orbitale triaxiale (van den Bosch & de Zeeuw 2010) et aux gradients de ratio de masse stellaire sur lumière (McConnell & Ma 2013), qui peuvent potentiellement affecter la détermination de la masse du SMBH.

### 1.4.3. Méthode directe : cinétique du disque mégamaser

La méthode par modélisation de la cinématique des disques mégamaser utilise l’imagerie de l’interférométrie à très longue base (VLBI – de l’anglais *Very Long Baseline Interferometry*) pour résoudre l’émission maser de la molécule d’eau,  $\text{H}_2\text{O}$  en rotation autour de SMBHs (Moriyama et al. 2019). Un maser est un processus par lequel des molécules amplifient les ondes électromagnétiques, spécifiquement dans le spectre des micro-ondes. Les mégamasers sont des masers extrêmement puissants associés aux SMBHs. Ceux-ci sont typiquement  $10^3 L_\odot$  plus lumineux qu’un maser régulier.

Le VLBI est basé sur le fait que dans les longueurs d’onde radio, il est possible de combiner le signal de plusieurs antennes ensemble pour former un télescope d’une taille équivalente à la plus grande distance entre deux antennes. C’est ce qu’on appelle interférométrie puisqu’il faut interférer les différents signaux ensemble en utilisant les composantes de Fourier de chaque signal.

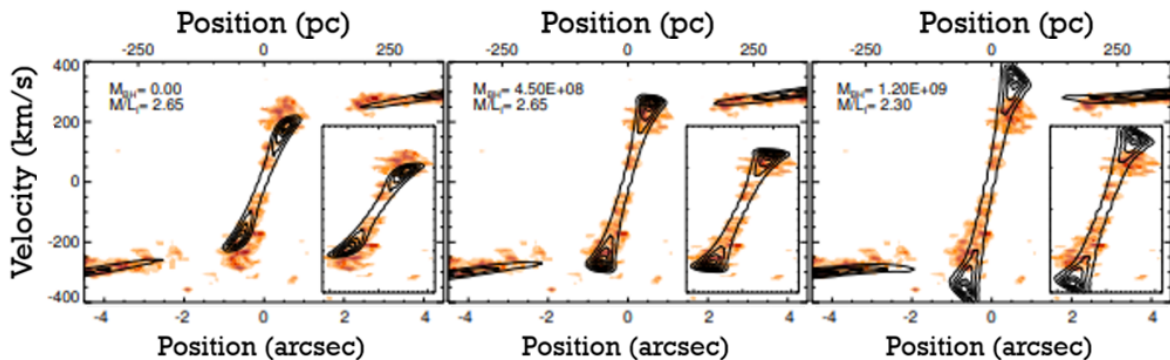
Ainsi, puisqu’un plus grand télescope permet une meilleure résolution spatiale, l’interférométrie est idéale pour mesurer la cinématique du gaz à des rayons très proches du SMBH et

donc obtenir sa masse. Ici, la masse du trou noir ne dépend que de la vitesse de rotation et du rayon du maser ( $M_{\text{BH}} = v^2 r / G$ ). Cette méthode est la méthode la plus précise pour mesurer la masse des SMBHs (Miyoshi et al. 1995; Kuo et al. 2011), puisqu'elle permet d'atteindre une résolution spatiale de seulement quelques milliarcsecondes. Par contre, ces disques sont habituellement trouvés uniquement dans certains types de galaxies spirales avec une masse de trous noirs faible, c'est-à-dire de  $\sim 10^7 M_{\odot}$ . Ce n'est donc pas une méthode idéale pour étudier les relations dans leur ensemble.

#### 1.4.4. Méthode directe : cinématique du gaz moléculaire

La méthode de la cinématique du gaz moléculaire froid est libre de presque tous les problèmes systématiques qui touchent habituellement les traceurs dynamiques chaud (étoiles et gaz ionisé) tout en utilisant des observations plus rapides et capables de résoudre amplement l'intérieur de la sphère d'influence. De plus, des études sur l'émission CO (Combes et al. 2007; Young et al. 2011; Alatalo et al. 2013) ont découvert que près de 10% de toutes les galaxies elliptiques proches contiennent une quantité significative de gaz froids moléculaires en leur centre en rotation régulière sous forme de disques. Plus de  $\sim 35000$  galaxies locales de tout type et grandeur pourraient être étudiées avec cette méthode (Davis et al. 2014). De plus, la dispersion de vitesse de ce gaz moléculaire possède une turbulence très faible ce qui réduit l'incertitude de la masse du SMBH mesurée (Alatalo et al. 2013). La première mesure de la masse d'un SMBH utilisant la cinématique du CO a été faite par Davis et al. 2013 et, depuis, l'Atacama Large Millimeter/submillimeter Array (ALMA) permet d'observer le CO à une résolution angulaire et une sensibilité supérieure (voir Figure 1.5, Barth et al. 2016c,b; Boizelle et al. 2017; Onishi et al. 2017; Davis et al. 2017, 2018b; Boizelle et al. 2019; Smith et al. 2019).

ALMA est un interféromètre situé au Chili composé de 66 antennes qui opèrent dans les longueurs d'onde submillimétrique soit de 31 à 1000 GHz. La distance minimale entre les antennes est de 150 m et celle maximale est de 16 km. ALMA possède 4 configurations (A, B, C, D) qui s'étendent de plus en plus. Une configuration compacte signifie une plus faible résolution spatiale, mais une meilleure sensibilité pour les variations de petites tailles (Schnee et al. 2014). Bien que l'interférométrie permet de simuler les effets d'une grande antenne, le processus de traitement des données qu'il faut réaliser pour obtenir une image dans le domaine spatial mène à certaines complications. En effet, un interféromètre ne prend pas une image directe du ciel, mais échantillonne l'espace des fréquences de la distribution de lumière du ciel qui est aussi nommé le plan U-V. La position des échantillons dans ce plan dépend de la distance entre chaque antenne. Un interféromètre immobile à 66 antennes échantillonnera  $66 \times 65 = 4290$  points dans le plan U-V, soit un point pour chaque combinaison de 2 antennes différentes. Dans notre cas, les antennes ne sont pas immobiles par rapport



**Figure 1.5** – Vitesse (km/s) dans la ligne de mire du gaz moléculaire de la galaxie NGC4526 en fonction de la distance radiale du trou noir supermassif au centre de celle-ci. Les points oranges correspondent aux données prises par ALMA et les courbes noires, l’ajustement pour 3 masses de trous noirs différentes : zéro (sous-estimation, à gauche),  $4.5 \times 10^8 M_\odot$  (le meilleur ajustement, au milieu) et  $1.2 \times 10^9 M_\odot$  (surestimation, à droite). Il est clair que la variation de la masse du trou noir n’affecte que la zone très proche de celui-ci, soit la sphère d’influence. Crédit : Davis et al. 2013.

aux objets observés considérant la rotation de la Terre. Normalement, pour des observations directes à haut temps d’exposition, cette rotation doit être corrigée, mais en interférométrie, elle déplace nos antennes, ce qui permet d’échantillonner un plus gros domaine de notre espace du plan U-V.

Idéalement, le traitement des données brutes acquises se ferait dans le domaine fréquentiel. Toutefois, puisqu’ici la mesure de la masse de SMBHs se fait par la corrélation spatiale de la cinématique du gaz moléculaire environnant, il est beaucoup plus simple d’analyser les données dans le domaine spatial. Cette transition du domaine fréquentiel au domaine spatial s’effectue par l’intermédiaire d’une transformation de Fourier inverse. Cependant, en raison d’un échantillonnage non exhaustif du plan U-V, l’image obtenue contient inévitablement des artefacts. Ces distorsions proviennent du fait que suite à l’échantillonnage non exhaustif du plan U-V, le signal dans le domaine spatial n’est plus entièrement localisé. Bien qu’elles renferment de l’information, ces artefacts compliquent le traitement des données. Un plus grand échantillonnage du plan U-V, que ce soit par l’ajout d’antennes ou par l’extension du temps d’exposition, atténue l’impact de ces artefacts. La correction de ces artefacts est souvent l’étape la plus longue en interférométrie parce qu’elle demande beaucoup de calcul informatique et une bonne compréhension des caractéristiques de l’objet observé pour pouvoir bien corrélérer les artefacts au vrai signal (CASA Team et al. 2022). En effet, certains artefacts, tels que les lobes secondaires, renferment de l’information, tandis que d’autres ne représentent que des nuisances.

Pour trouver la masse de trous noirs avec la méthode de la modélisation de la cinématique du gaz moléculaire, il suffit d’utiliser la même technique que décrite dans la Section 1.4.1

pour le gaz ionisé, mais cette fois d'utiliser la ligne d'émission  $(2 - 1)\text{CO}$  ou  $(3 - 2)\text{CO}$  qui est beaucoup plus froide et sans extinction. En astronomie, l'extinction fait référence à l'absorption et à la dispersion de la lumière par la poussière et le gaz situés entre l'objet observé et l'observateur. De plus, contrairement à HST, pour lequel il faut prendre plusieurs fentes d'observations, les observations avec ALMA se prennent directement sous la forme d'un cube de données (2 dimensions spatiales et une dimension spectrale). Il est donc possible d'avoir toutes les données nécessaires en une seule observation. Cette méthode est si efficace, précise et exacte qu'une récente collaboration établie en 2017 nommée WISDOM (*mm-Wave Interferometric Survey of Dark Object Masses*) a comme objectif d'utiliser ALMA pour mesurer un maximum de SMBH dans le but d'éclaircir notre compréhension de la co-évolution entre SMBH et leur galaxie hôte (Onishi et al. 2017; Davis et al. 2017, 2018b; Smith et al. 2019; North et al. 2019; Davis et al. 2020; Smith et al. 2021a,b; Lelli et al. 2022).

### 1.4.5. Méthode indirecte : plan fondamental de l'activité des trous noirs

Une autre façon de trouver la masse d'un trou noir, cette fois-ci par méthode indirecte, est d'utiliser le plan fondamental de l'activité des trous noirs (e.g. Merloni et al. 2003; Falcke et al. 2004; KÖrding et al. 2006; Gültekin 2009; Plotkin et al. 2012; Hlavacek-Larrondo et al. 2012a; Mezcua et al. 2018b), qui corrèle la masse, la luminosité rayons X ( $\log L_{X\text{-ray}}$ ) qui provient du disque d'accrétion et la luminosité radio ( $\log L_{\text{radio}}$ ) de certains trous noirs par l'équation :

$$\log M_{\text{BH}} = a \log L_{X\text{-ray}} + b \log L_{\text{radio}} + c. \quad (1.4.2)$$

Cette relation est vraie pour les trous noirs de toutes masses, des trous noirs stellaires ( $3 M_{\odot}$ ) aux SMBHs ( $\sim 10^{10} M_{\odot}$ ). La relation considérée comme la plus robuste jusqu'à présent a été trouvée par Plotkin et al. 2012 et donne des valeurs de coefficients de  $a = 1,14$ ,  $b = 1,65$  et  $c = -6,90$ . La luminosité radio a été trouvée à 5 GHz et la luminosité rayons X dans la bande de  $0,5 - 10$  keV. Les points noirs dans la Figure 1.6 représentent des systèmes où les trois valeurs étaient connues. Il serait donc possible d'utiliser l'Équation 1.4.2 pour inférer la masse de trous noirs dont les luminosités sont connues.

Le plan fondamental de l'activité des trous noirs permet aussi d'étudier les deux relations de la Figure 1.4,  $M_{\text{BH}} - \sigma_{\star}$  et  $M_{\text{BH}} - L_{K,\text{bulbe}}$ , et voir si, en effet, elles ne tiennent plus aux extrémités de la distribution. Dans la section 1.3, il a été mentionné que les SMBHs d'une masse de plus de  $10^{10} M_{\odot}$  ne semblent pas suivre les relations  $M_{\text{BH}} - \sigma_{\star}$  et  $M_{\text{BH}} - L_{K,\text{bulbe}}$ . Des

études (Hlavacek-Larrondo et al. 2012b; Mezcua et al. 2018a) ont examiné le positionnement de ces trous noirs dans le plan fondamental si la masse des trous noirs était déterminée à l'aide des relations  $M_{\text{BH}} - \sigma_*$  et  $M_{\text{BH}} - L_{K,\text{bulbe}}$ . Mezcua et al. 2018a a fait cette étude pour 72 galaxies dans l'Univers local, voir Figure 1.6, où les points bleus et rouges représentent des trous noirs dans des galaxies dont la masse a été calculée avec la relation  $M_{\text{BH}} - M_{K,\text{bulbe}}$  ( $L_{K,\text{bulbe}}$  étant relié à  $M_{K,\text{bulbe}}$ ). Basé sur le plan fondamental, il semblerait que cette relation sous-estime les masses d'un facteur  $\sim 10$  pour les SMBHs dans les galaxies dominantes, soit celles qui se trouvent au centre d'amas de galaxies (Mezcua et al. 2018a). Selon cette étude,  $\sim 40\%$  de ces SMBHs devaient avoir une masse de plus de  $10^{10} M_{\odot}$  et donc les relations locales entre les masses des trous noirs et les propriétés des galaxies ne tiendraient pas dans les galaxies dominantes. Ceci pourrait être dû à une formation en deux phases, le SMBH en premier, puis la galaxie, ou ces SMBHs pourraient provenir de trous noirs formés à un haut décalage vers le rouge. Ceci semble donc aller à l'encontre de la théorie de l'évolution conjointe des trous noirs et de leur galaxie hôte. C'est pour cela qu'il est important de mesurer plus de masses de trous noirs dans différents types de galaxies pour confirmer ou infirmer ces études.

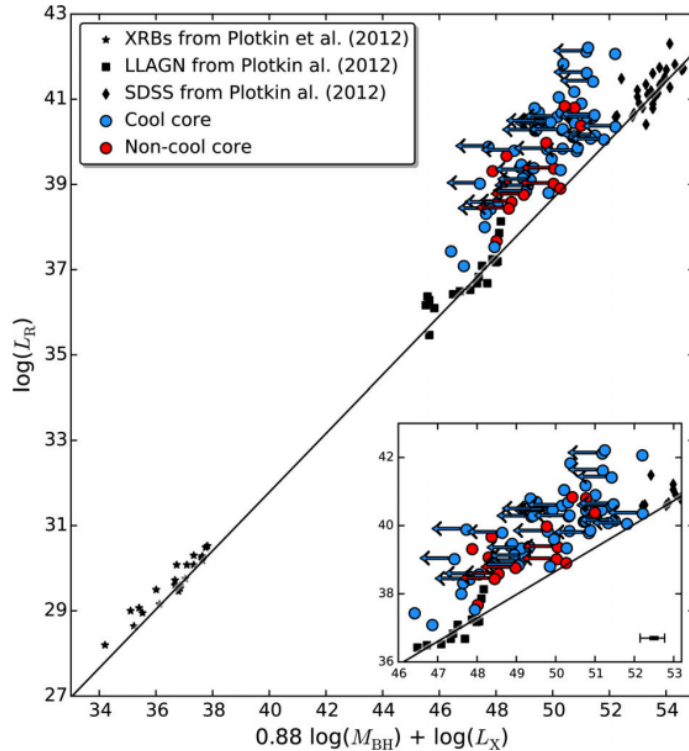
#### 1.4.6. Méthode indirecte : profil de brillance de surface

Une autre méthode indirecte pour trouver la masse de trous noirs est d'utiliser le profil de brillance de surface stellaire de la galaxie hôte. Certaines galaxies qui contiennent un SMBH très massif en leur centre vont avoir ce qu'on appelle un cœur ou un déficit de lumière dans leur profil de brillance. La théorie dominante pour expliquer ce déficit de lumière propose qu'il soit produit par deux trous noirs binaires qui expulsent des étoiles de la galaxie due à leur mouvement (Begelman et al. 1980). Par contre, ce qui est particulièrement important pour l'étude des SMBHs, c'est que la taille de ce déficit de lumière dans le profil de brillance corrèle avec la masse totale d'étoiles éjectées et donc avec la masse du SMBH central (Lauer et al. 2007b).

#### 1.4.7. Méthode indirecte : cartographie par réverbération

La cartographie par réverbération est une technique puissante utilisée en astrophysique pour mesurer la masse de SMBHs. Lorsqu'un SMBH absorbe activement de la matière de son environnement, il crée un disque d'accrétion composé de gaz et de poussière en rotation qui s'enfonce progressivement vers le SMBH. Ce disque de matière tourne autour du SMBH à une vitesse si élevée que les collisions entre particules amplifient l'énergie thermique du disque, à tel point qu'il émet des photons sur un large spectre de longueurs d'onde, illuminant de ce fait le gaz situé à de grandes distances (Peterson et al. 2004). Cette zone éloignée où





**Figure 1.6** – Le plan fondamental de l’activité des trous noirs représentant la relation entre la luminosité radio ( $L_{\text{radio}}$ ) et rayons X ( $L_{\text{X-ray}}$ ) et la masse des trous noirs (ligne noire). Les symboles noirs proviennent de Plotkin et al. 2012 et représentent différents types de trous noirs. Les points en rouge et en bleu représentent des SMBHs au centre de la galaxie dominante d’amas de galaxies dont la masse a été mesurée avec la relation  $M_{\text{BH}} - M_{K,\text{bulbe}}$  en connaissant la valeur de  $M_{K,\text{bulbe}}$ . La sous-estimation de la masse de SMBHs par le plan fondamental est clairement visible. Crédit : Mezcua et al. 2018a.

le gaz ionisé se trouve est appelée la région à larges raies (BLR – de l’anglais "Broad-Line Region").

La cartographie par réverbération exploite le délai temporel entre les variations de luminosité du disque d’accrétion et la réponse correspondante des raies d’émission de la BLR. Étant donné que le disque d’accrétion est plus proche du SMBH que la BLR, les changements de luminosité du disque sont détectés en premier (Vestergaard & Peterson 2006). Ce délai ( $\tau$ ) est directement lié à la masse du SMBH :

$$M_{\text{SMBH}} = \frac{f c \tau v^2}{G}, \quad (1.4.3)$$

où  $v$ , la vitesse du gaz dans la BLR et  $f$ , un facteur d'échelle sans dimension inféré à partir d'autres observations qui encode la structure, la cinématique et l'orientation de la BLR (Onken et al. 2004). L'Équation 1.4.3 établit simplement une relation entre la masse de l'objet central, la vitesse et le rayon ( $R_{BLR}$ ) d'une masse en orbite autour de celui-ci. Dans ce contexte, la vitesse peut être directement obtenue par spectroscopie (Bentz et al. 2013) et le rayon est déterminé à partir du délai temporel :  $R_{BLR} = c\tau$  (Peterson 1993).

## 1.5. Les simulations cosmologiques

En astrophysique, les simulations jouent un rôle crucial, à la fois pour la validation des théories en reproduisant des phénomènes observés et pour l'étude de phénomènes actuellement non observables. Il existe plusieurs méthodes de simulation, notamment les techniques analytiques, les méthodes semi-analytiques et les simulations numériques. Bien que les techniques analytiques soient relativement simples à implémenter, leur utilisation reste très limitée puisque l'équation évolutive du système doit être connue. Les méthodes semi-analytiques, quant à elles, sont plus expressives, mais nécessitent l'établissement de nombreux paramètres libres pour reproduire fidèlement les observations. En revanche, les simulations numériques, en particulier les simulations à  $N$  corps et hydrodynamiques, sont largement privilégiées en raison de leur correspondance avec les observations (Navarro et al. 1997). Néanmoins, elles sont très exigeantes d'un point de vue computationnel en raison de la multitude de processus physiques à modéliser : la croissance des halos de matière noire par fusion, la dynamique du refroidissement du gaz (Bullock et al. 2001), la transformation du gaz froid en étoiles, l'évolution spectrophotométrique et chimique de la population stellaire, la rétroaction énergétique des étoiles (Valiev et al. 2010), l'évolution des propriétés du milieu interstellaire (ISM – de l'anglais *interstellar medium*), et la formation de la galaxie à travers des événements de fusion successifs (Bower et al. 2006) et la rétroaction de trous noirs.

Avant de commencer la génération d'une simulation numérique, il est nécessaire de définir la taille et la cosmologie de l'Univers dans lequel tout évoluera. Cette cosmologie régit l'expansion de l'Univers ainsi que la distribution de la matière et de l'énergie en fonction du temps,  $t$ . Elle est gouvernée par l'équation de Friedmann 1922 :

$$\frac{H(t)^2}{H_0^2} = \Omega_{0,R}a(t)^{-4} + \Omega_{0,M}a(t)^{-3} + \Omega_{0,k}a(t)^{-2} + \Omega_{0,\Lambda}, \quad (1.5.1)$$

où  $H_0$  représente la constante de Hubble,  $\Omega_{0,R}$  la densité de radiation (photons) actuelle,  $\Omega_{0,M}$  la densité de matière (noire et baryonique) actuelle,  $\Omega_{0,k}$  la densité de la courbure spatiale,  $\Omega_{0,\Lambda}$  la densité du vide (la constante cosmologique), et  $a(t)$  un facteur d'échelle.

Cette équation est dérivée du modèle de matière noire froide avec constante cosmologique ( $\lambda$ CDM – de l’anglais *lambda cold dark matter*, Planck Collaboration et al. (2016)).

### 1.5.1. Les simulations à N corps

La matière noire, de nature non collisionnelle, n’interagit avec son environnement que par l’attraction gravitationnelle, ce qui la rend particulièrement propice aux simulations à N corps. Ces simulations se caractérisent par l’interaction et l’évolution de N éléments indépendants, dans ce cas, l’élément est la matière noire. La distribution initiale de la densité de matière sombre, notée  $f$ , dans la simulation est déterminée par le spectre de puissance du modèle cosmologique sélectionné, dans ce cas le  $\lambda$ CDM. L’évolution temporelle de cette densité est régulée par l’équation de Boltzmann (Boltzmann 1894) non collisionnelle :

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \vec{v} \cdot \nabla f - \frac{\partial f}{\partial \vec{v}} \cdot \nabla \Phi, \quad (1.5.2)$$

où  $\vec{v}$  représente la vitesse et  $\Phi$  symbolise le potentiel gravitationnel, ce dernier étant décrit par l’équation de Poisson (Poisson 1824).

Toutes les particules de matière noire ont la même masse,  $M_s$ , et c’est elle qui définit la résolution de la simulation. Leur unique interaction se manifeste par la force d’attraction gravitationnelle. Pour calculer la force gravitationnelle totale,  $F_i$ , qui agit sur une particule,  $i$ , de matière sombre, il faut prendre en compte la somme de toutes les interactions avec les N-1 autres particules. À chaque intervalle de temps discret,  $\Delta t$ , la simulation à N corps doit résoudre le système d’équations suivant, composé de trois équations :

$$\vec{F}_i(t) = -Gm_s \sum_{j \neq i}^{N-1} \frac{\vec{r}_i(t) - \vec{r}_j(t)}{(|\vec{r}_i(t) - \vec{r}_j(t)|^2 + \epsilon^2)^{3/2}}, \quad (1.5.3)$$

$$\vec{v}_i(t + \Delta t) = \vec{v}_i(t) + \frac{F_i}{m_s} \Delta t, \quad (1.5.4)$$

$$\vec{r}_i(t + \Delta t) = \vec{r}_i(t) + \vec{v}_i \Delta t, \quad (1.5.5)$$

ici,  $j$  est l'indice représentant toutes les autres particules de matière noire,  $\vec{r}$  désigne le vecteur de position tridimensionnel et  $\epsilon$  est une longueur d'atténuation. Cette dernière permet d'éliminer la singularité de l'équation lorsque  $\vec{r}_i = \vec{r}_j$  et de stabiliser numériquement les situations où  $\vec{r}_i \approx \vec{r}_j$ . Il est crucial de noter que ce système d'équations est exprimé dans le système de coordonnées comobiles, c'est-à-dire que l'expansion de l'Univers est prise en compte de telle sorte que la distance entre deux objets statiques reste constante avec le temps.

L'évolution temporelle du système à N corps est modélisée en répétant, à chaque pas de temps (de taille variable), les opérations suivantes. Tout d'abord, une double itération est nécessaire pour calculer la force agissant sur chaque particule à l'aide de l'équation 1.5.3. Ensuite, une boucle permet de résoudre l'équation différentielle pour la vitesse (voir l'Équation 1.5.4), où la nouvelle vitesse est obtenue en ajoutant à la vitesse précédente le produit de l'accélération par le pas de temps. Enfin, une autre itération permet de résoudre l'équation différentielle pour les positions (voir l'Équation 1.5.5) : la nouvelle position est déduite de l'ancienne en y ajoutant le produit de la vitesse par le pas de temps. Le compteur de temps est ensuite mis à jour et toutes les étapes précédentes sont répétées à partir des nouvelles variables. À chaque pas de temps, il est donc nécessaire de mémoriser la position  $\vec{r} = (r_x, r_y, r_z)$  et la vitesse  $\vec{v} = (v_x, v_y, v_z)$  de chaque particule. Pour les N corps, il faut donc garder en mémoire  $6N + 1$  paramètres, le paramètre supplémentaire étant requis pour suivre l'évolution temporelle. Cette méthode de résolution directe nommé particule par particule (PP – de l'anglais *Particle-Particle*) est certes simple, mais elle est également extrêmement coûteuse en termes de calcul (Kobayashi 2004). En effet, une complexité de  $N^2$  est nécessaire pour calculer la force entre chaque particule. Cette complexité provient du besoin de calculer toutes les distances entre toutes les particules à chaque pas de temps. Bien que cette méthode soit exacte, elle devient impraticable lorsque le nombre N de corps devient trop important. Par exemple, une simulation impliquant 100,000 particules nécessiterait un peu moins d'une heure de calcul par pas de temps.

La méthode de particule par maillage (PM – de l'anglais *Particle-Mesh*) résout le système d'équations à chaque pas de temps en regroupant les particules individuelles au sein d'un maillage de résolution arbitraire (Klypin & Holtzman 1997). Par interpolation, les positions des particules sont utilisées pour déterminer le potentiel gravitationnelle causé par la masse dans chaque cellule du maillage, puis la force entre chaque cellule est calculée. Le temps de calcul est nettement réduit car, en fonction de la résolution choisie, le nombre de cellules sera beaucoup plus faible que le nombre de particules. Une fois la force agissant sur chaque cellule déterminée, le même processus d'interpolation permet d'obtenir la force qui agit sur chaque particule. Une fois  $F_i$  connue, les Équations 1.5.4 et 1.5.5 sont facilement résolues, car la vitesse et la position de chaque particule sont constamment mémorisées. La méthode PM réduit considérablement la complexité du calcul de la force en regroupant les particules

dans un maillage, accélérant ainsi la simulation et permettant son application à des systèmes comportant un très grand nombre de particules. Cependant, les résultats obtenus sont approximatifs. C'est à petite échelle que cette imprécision est la plus marquée. Idéalement, une méthode qui traite les particules comme un ensemble à grande échelle et comme des entités indépendantes à petite échelle serait privilégiée.

Il est possible d'atténuer le problème d'inexactitude en introduisant un maillage adaptatif, dont la résolution varie en fonction de la densité des particules : plus une zone est dense, plus la résolution est élevée, permettant de prendre en compte les effets à petite échelle (Theuns et al. 1998). Cependant, cette approche ne résout pas entièrement les problèmes de résolution introduits par le maillage, elle les atténue simplement. À cet égard, la méthode particule par particule et particule par maillage (P3M, de l'anglais *particle-particle/particle-mesh*) est souvent préférée. Cette dernière divise la force gravitationnelle en deux composantes : la composante à courte portée est calculée en utilisant une PP, tandis que la composante à longue portée est obtenue par une méthode PM (Harnois-Déraps et al. 2013). Ces deux composantes nécessitent une liaison cohérente définie par une distance critique arbitraire. Si cette distance est faible, la méthode P3M se rapproche d'une méthode PM ; si elle est grande, elle tend vers une méthode PP.

Le calcul de la force est effectué par sommation directe pour les particules proches, tandis que pour les particules plus éloignées, les propriétés des groupes de particules sont exploitées, réduisant considérablement le nombre de distances à estimer. Cependant, il est nécessaire de construire et de stocker de manière adéquate les informations sur les groupes de particules. Cela est réalisé en utilisant la technique des arbres (Hernquist 1987). En effet, l'algorithme d'arbre de Barnes-Hut est l'un des plus populaires (Rein & Liu 2012). Il génère une hiérarchie de cellules en considérant initialement la simulation comme une unique cellule, qui est ensuite subdivisée en un nombre arbitraire et constant de sous-cellules. Si une sous-cellule contient plus d'une particule, elle est à nouveau subdivisée. Si elle contient une seule particule, le processus s'arrête ; si elle est vide, elle est ignorée. Cet arbre doit être généré à chaque pas de temps.

Pour utiliser l'arbre et choisir entre les méthodes PP et PM, une distance critique doit être établie (Deserno & Holm 1998). En pratique, une fois l'arbre construit, pour une particule donnée, il faut considérer le groupe de particules le plus grand de l'arbre, à savoir celui situé au sommet. Si la distance entre la particule et le groupe observé est inférieure à la distance critique, la contribution gravitationnelle est considérée dans son ensemble. Sinon, on descend dans la hiérarchie des nœuds et on répète le processus. Cette étape est appelée la descente de l'arbre. Bien que la méthode P3M nécessite davantage de mémoire, car elle doit conserver l'arbre, elle offre le meilleur des deux mondes en termes de simulation à N corps PP et PM (MacFarland et al. 1998).

Plusieurs tests peuvent être effectués pour vérifier la validité d’une simulation à  $N$  corps basés sur le principe de conservation. En effet, la simulation doit veiller à conserver la masse, l’énergie, l’impulsion et le moment angulaire du système à chaque pas de temps. La masse est facilement conservée, car le nombre de particules ne change pas. Avec la méthode P3M, il est relativement aisé d’obtenir une conservation de l’énergie et de l’impulsion avec moins de 0.1% d’écart (Dubinski et al. 2004). C’est la conservation du moment angulaire qui peut poser le plus de problèmes, notamment en raison des conditions limites qui apparaissent lorsqu’une méthode faisant appel au maillage est utilisée (Chen et al. 2015).

### 1.5.2. Les simulations hydrodynamiques

La matière noire est sans collision, c’est-à-dire que sa section efficace de collision est négligeable. En conséquence, la seule force à l’œuvre est la gravité. La matière baryonique, en revanche, interagit à la fois par gravité et par divers processus radiatifs : effet Compton, Bremsstrahlung, ionisation, etc. Pour simuler la matière baryonique, il est donc nécessaire d’intégrer une simulation hydrodynamique qui évolue simultanément à la simulation à  $N$  corps.

En physique des fluides, les équations typiques sont tridimensionnelles et formulées sous forme d’équations de conservation du flux. L’équation d’Euler 1936 régit la dynamique du fluide :

$$\frac{d\vec{u}}{dt} = -\nabla\Phi - \frac{\nabla P}{\rho}, \quad (1.5.6)$$

où  $\rho$  est la densité,  $\vec{u}$  le débit du fluide et  $P$  l’impulsion. Pour une simulation hydrodynamique, le fluide doit respecter la loi de conservation de la continuité. L’équation de continuité, qui doit être résolue à chaque instant en fonction de  $\rho$  et  $\vec{u}$ , garantit la conservation du flux d’impulsion du fluide :

$$\frac{d\rho}{dt} + \nabla \cdot (\rho\vec{u}) = 0. \quad (1.5.7)$$

En général, pour résoudre les équations hydrodynamiques, deux types de techniques numériques sont couramment utilisées : les méthodes eulériennes et les méthodes lagrangiennes. Les méthodes eulériennes consistent à calculer les valeurs moyennes des paramètres du fluide en fonction d’un maillage à résolution fixe ou adaptative (Springel 2005). De ce fait, les

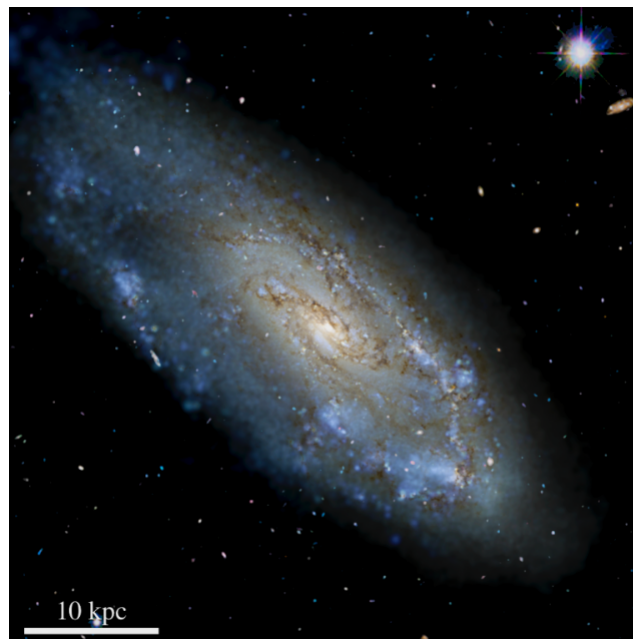
équations de conservation sont résolues en fonction des distances prédéfinies entre les cellules du maillage. À chaque pas de temps, il est nécessaire de calculer les valeurs moyennes pour chaque cellule, d'évaluer l'interaction d'une cellule spécifique avec son environnement en fonction des Équations 1.5.6 et 1.5.7, et d'actualiser les valeurs moyennes une fois que le flux de chaque cellule a été déterminé. Le flux d'une cellule  $i$  quelconque dépend uniquement de la matière dans les cellules directement adjacentes, soit  $i \pm 1$ . L'avantage de cette méthode est que le calcul du gradient entre chaque cellule est simple et précis (Schaye et al. 2015). Toutefois, comme mentionné précédemment pour les simulations à N corps, l'utilisation d'un maillage peut introduire des problèmes de résolution et de gestion de la mémoire.

Les méthodes lagrangiennes, qui considèrent les paramètres de chaque particule individuellement, permettent d'éviter l'utilisation d'un maillage. Une méthode lagrangienne couramment utilisée est l'hydrodynamique des particules lissées (SPH – de l'anglais *Smoothed Particle Hydrodynamics*, Springel & Hernquist 2003). La SPH calcule les propriétés du fluide en tout point en se basant sur une moyenne pondérée des particules voisines. Comme les interactions du fluide sont locales, seules les particules proches sont prises en compte dans le calcul de la moyenne. La résolution de la simulation devient alors directement proportionnelle à la densité : plus il y a de particules à un endroit donné, plus la moyenne sera sensible à un changement de position (Hernquist 1987). Il devient ainsi possible d'inférer les propriétés du fluide sans nécessité de maillage. Bien que les méthodes eulériennes soient limitées en termes de résolution spatiale, leur résolution en masse est élevée, car toutes les particules sont prises en compte dans le maillage (Kereš et al. 2005). Pour les méthodes lagrangiennes, la résolution spatiale est très élevée, mais la résolution en masse est plus faible puisque c'est la moyenne des particules qui est considérée (Stinson et al. 2006). Le choix de la méthode employée dépend du type de simulation et du processus physique que l'on cherche à étudier.

La majorité des suites de simulations cosmologiques actuelles telles que IllustrisTNG (Pillepich et al. 2018), EAGLE (Crain et al. 2015) et Simba (Davé et al. 2019) simulent l'évolution de plusieurs centaines de milliers de galaxies. Bien qu'elles soient idéales pour étudier les relations à grande échelle, elles ne sont pas suffisamment précises pour analyser les effets à petite échelle des SMBHs sur leur environnement, ce qui est l'objectif de notre projet. À cet égard, la suite de simulations hydrodynamiques FIRE (de l'anglais *Feedback In Realistic Environments*) s'avère être la simulation la plus adaptée pour mesurer de façon directe la masse de SMBHs.

FIRE est une simulation cosmologique axée sur l'évolution temporelle d'une seule galaxie, conçue pour étudier sa formation et son évolution (Hopkins et al. 2018). Dans cette suite, les propriétés de la matière noire sont simulées en utilisant une simulation à N corps et la méthode P3M. Pour la matière baryonique, elle est traitée avec GIZMO, un algorithme qui allie des méthodes eulériennes et lagrangiennes. Un maillage adaptatif est introduit dans la simulation et ajusté par la méthode SPH (Wetzel et al. 2023). Les simulations FIRE intègrent

également des modèles sous-grille, prenant en compte les effets des processus physiques cruciaux se produisant à des échelles inférieures à la résolution de la simulation. Ces processus comprennent la formation stellaire, le refroidissement radiatif, la rétroaction stellaire et les effets des trous noirs sur leur environnement (Ma et al. 2016). Ces modèles sous-grille sont calibrés en utilisant des simulations de haute résolution à petite échelle et sont soigneusement ajustés pour reproduire les propriétés observées des galaxies dans l’Univers local (Grudić et al. 2023). En se concentrant sur l’évolution d’une seule galaxie, les effets à très petite échelle peuvent être codés et conservés en mémoire. Pour étudier les effets d’un SMBH sur son environnement, une précision à petite échelle est indispensable, ce qui rend FIRE particulièrement adapté à cette tâche.



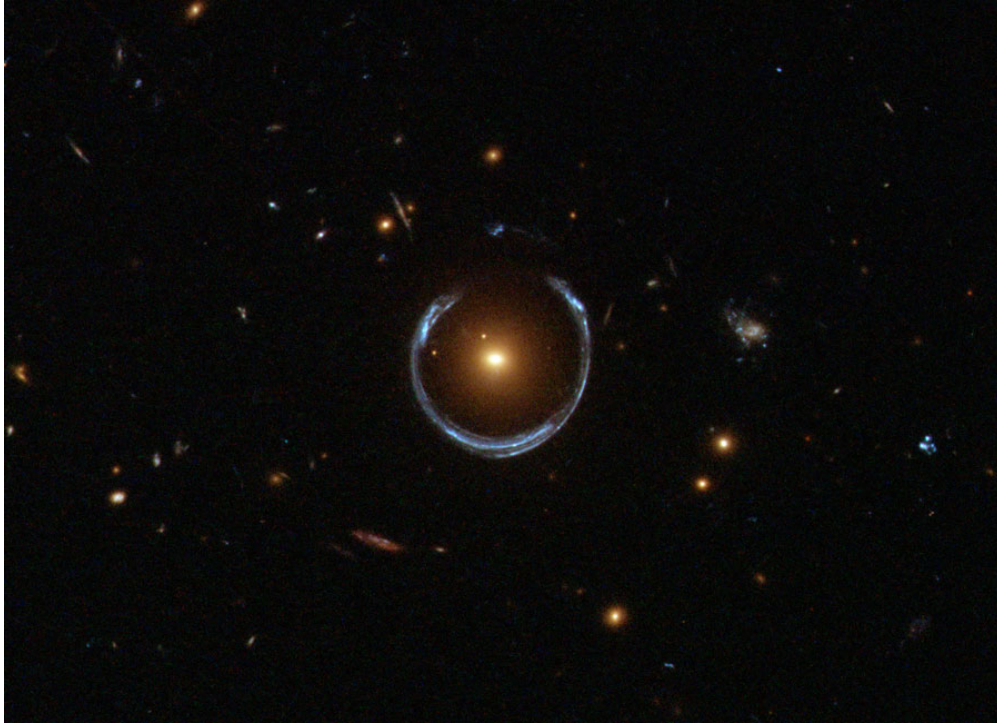
**Figure 1.7** – Image d’une galaxie de type Voie Lactée générée par la simulation cosmologique FIRE tel qu’observée par HST. Crédit : Hopkins 2015.

## 1.6. Les lentilles gravitationnelles

En considérant la relativité générale d’Einstein (Einstein 1936), nous savons que la masse engendre une courbure de l’espace-temps affectant ainsi le déplacement de toute matière dans cet environnement. De ce fait, comme le soleil qui influe sur le mouvement de la Terre, il est possible de dévier la trajectoire de la lumière. Bien que la lumière n’ait pas de masse, elle ne peut échapper à la courbure de l’espace-temps dans lequel elle se propage. C’est de ce principe que naissent les lentilles gravitationnelles. En bref, une lentille gravitationnelle se produit lorsqu’un corps céleste très massif (ex. : un amas de galaxies) dévie les rayons lumineux émis par un objet situé derrière lui, déformant ainsi l’image observée. Cette déformation se traduit



généralement par la formation d'un ou plusieurs arcs lumineux autour de l'objet en premier plan (voir Figure 1.8). Par ailleurs, l'effet d'une lentille gravitationnelle est achromatique, ce qui signifie que tous les photons, quel que soit leur spectre de fréquence, seront déviés de la même manière. En effet, la courbure de l'espace-temps ne varie pas en fonction de la longueur d'onde de la matière qui s'y propage (Einstein et al. 1938).



**Figure 1.8** – Image d'une lentille gravitationnelle prise par le télescope spatial Hubble d'une galaxie massive en avant plan qui déforme l'image observée d'une galaxie en arrière plan. L'alignement entre ces galaxies est si précis que la galaxie en arrière-plan est déformée en un fer à cheval, formant ainsi un anneau presque complet. Crédit : ESA/Hubble & NASA.

Les déformations causées par les lentilles gravitationnelles peuvent, dans certains cas, nous permettre d'étudier des objets à des distances très éloignées. En effet, les rayons lumineux de certains de ces arcs peuvent être focalisés, étirés et/ou agrandis, ce qui nous permet d'étudier des objets peu lumineux et petits à des distances qui nous seraient normalement inaccessibles. Toutefois, pour pouvoir exploiter ces déformations, la reconstitution précise de l'image en arrière-plan est essentielle, une procédure qui nécessite l'application de l'équation de lentille.

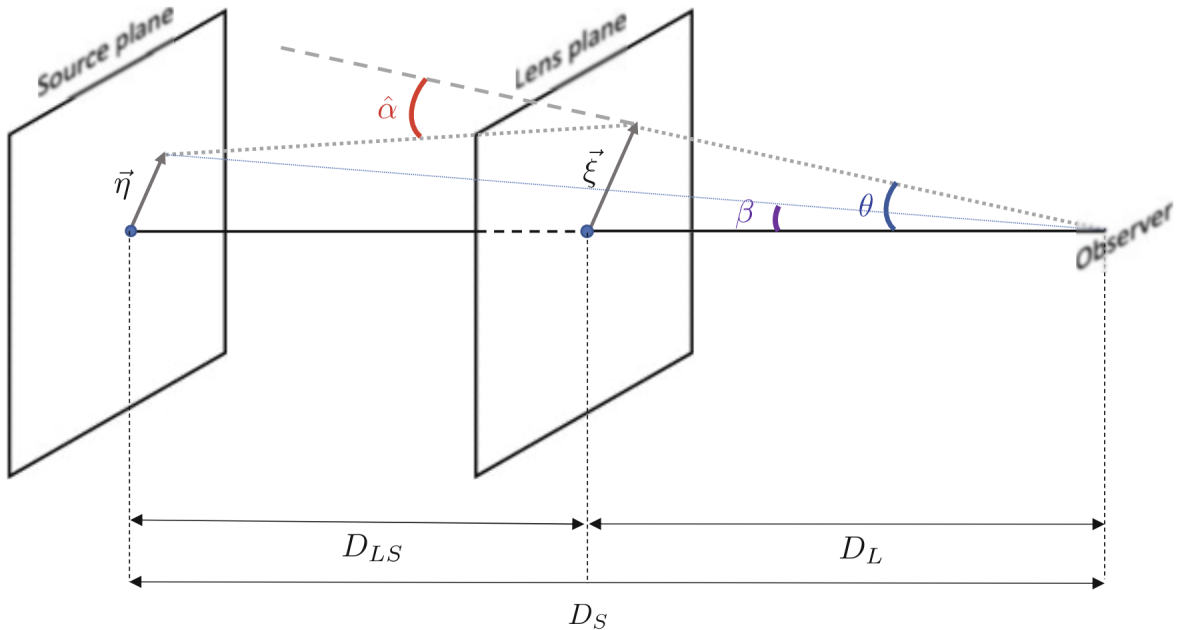
### 1.6.1. L'équation de lentille

Tout d'abord, il faut définir la relation entre la position apparente et intrinsèque de l'objet situé en arrière-plan. En l'absence d'objet massif en premier plan, l'objet est observé à une position particulière dans le ciel, appelée position intrinsèque de la source  $\beta$  (en unité

angulaire). Toutefois, si un effet de lentille gravitationnelle est présent, les rayons lumineux seront déviés, et la source sera observée à une nouvelle position  $\theta$ , correspondant à sa position apparente (Meneghetti 2022).

Pour bien comprendre la situation, nous nous baserons sur le schéma (voir Figure 1.9) d'une lentille gravitationnelle typique. La masse source est placée à un décalage vers le rouge  $z_S$  (de l'anglais *redshift*) correspondant à une distance angulaire  $D_S$ . La lentille est, elle, placée à  $z_L$ , soit une distance angulaire  $D_L$ . Ainsi, il est possible de définir  $D_{SL}$ , la distance angulaire entre le plan de la source et celui de la lentille. À noter qu'une distance angulaire est définie comme le rapport entre la taille physique transversale de l'objet et sa taille angulaire. De ce fait, il est possible de passer de la séparation physique à la séparation angulaire par rapport à la distance angulaire. Cependant, puisque l'expansion de l'univers varie en fonction de la distance (constante de Hubble,  $H_0$ ), alors

$$D_{LS}(z_S, z_L) \neq D_S(z_S) - D_L(z_L). \quad (1.6.1)$$



**Figure 1.9** – Représentation schématique de la distorsion d'un faisceau lumineux causé par une lentille gravitationnelle. La source est placée à une distance angulaire  $D_S$  de l'observateur et la lentille, à  $D_L$ . L'angle de déflexion  $\alpha$  est le paramètre communément utilisé pour quantifier la déviation du faisceau. Crédit : Meneghetti 2022.

Avant de se lancer dans le développement mathématique de l'équation de lentille, il est important de considérer l'approximation du plan mince. L'idée ici est que, puisque les distances  $D_S$ ,  $D_L$  et  $D_{LS}$  sont beaucoup plus grandes que la taille physique de la source ainsi que de la lentille, il est possible de négliger l'extension dans la ligne de visée de la source et de la lentille lors du calcul de la déflexion des rayons lumineux. En d'autres termes, la source et la lentille peuvent être considérées comme deux plans distincts infiniment mince à une distance respective de  $D_S$  et  $D_L$ . Notons que puisque l'image apparente de la source provient du plan lentille, celui-ci est souvent appelé plan image.

Pour bien comprendre la corrélation entre la position intrinsèque et la position apparente, nous allons suivre la trajectoire d'un faisceau lumineux émis par la source, dévié par la lentille et arrivant à l'observateur. Ce faisceau est représenté par la ligne pointillée sur la Figure 1.9.

Sachant que la position angulaire vectorielle intrinsèque de la source est  $\vec{\beta}$ , nous pouvons déduire que la distance physique au plan source est  $\vec{\eta} = D_S \vec{\beta}$ . De même, pour une position apparente de  $\vec{\theta}$ , la distance physique au plan image est  $\vec{\xi} = D_L \vec{\theta}$  et celle au plan source est  $D_S \vec{\theta}$ . Ainsi, en définissant  $\hat{\alpha}(\vec{\theta})$  comme étant l'angle de déviation du faisceau lumineux,  $D_{LS} \hat{\alpha}(\vec{\theta})$  représente le décalage physique dû à l'effet de lentille au niveau du plan source. De ces constats découle l'équation de lentille définie au plan source.

$$D_S \vec{\theta} = D_S \vec{\beta} + D_{LS} \hat{\alpha}(\vec{\theta}). \quad (1.6.2)$$

En définissant un angle de déviation réduit, soit

$$\vec{\alpha}(\vec{\theta}) = \frac{D_{LS}}{D_S} \hat{\alpha}(\vec{\theta}), \quad (1.6.3)$$

on obtient

$$\vec{\beta} = \vec{\theta} - \vec{\alpha}(\vec{\theta}). \quad (1.6.4)$$

Il est important de noter que la fonction  $\vec{\alpha}(\vec{\theta})$  est une fonction de la position apparente  $\vec{\theta}$ , car l'angle de déviation dépend de la position du faisceau lumineux au plan image. En effet, la trajectoire de la lumière émise par la source est déviée par la courbure gravitationnelle de

la lentille. Cette courbure n'est bien évidemment pas identique en tout point, mais est une fonction du potentiel gravitationnel de la lentille qui est directement proportionnelle à la masse (Li & Mandelbaum 2023). Ainsi, étant donné que la courbure de l'espace-temps due à la lentille varie en fonction de la position au plan lentille, la déviation des photons émis par la source  $\vec{\alpha}$  dépend de  $\vec{\theta}$ .

Il est commun et utile de redéfinir l'Équation 1.6.4 sous forme adimensionnelle. Pour cela, on commence par définir une distance d'échelle dans le plan image  $\xi_0$ , qui correspond à une distance d'échelle au plan source  $\eta_0 = \xi_0 D_S / D_L$ . Avec ces nouvelles variables, nous définissons les vecteurs adimensionnels et l'angle de déviation d'échelle suivants

$$\vec{x} \equiv \frac{D_L \vec{\theta}}{\xi_0} = \frac{\vec{\xi}}{\xi_0}, \quad (1.6.5)$$

$$\vec{y} \equiv \frac{D_S \vec{\beta}}{\eta_0} = \frac{\vec{\eta}}{\eta_0}, \quad (1.6.6)$$

$$\vec{\alpha}(\vec{x}) = \frac{D_L D_{LS}}{\xi_0 D_S} \hat{\alpha}(\vec{\theta}) = \frac{D_L}{\xi_0} \vec{\alpha}(\vec{\theta}). \quad (1.6.7)$$

En remplaçant ces trois relations adimensionnelles au sein de l'Équation 1.6.4, on obtient l'équation de lentille adimensionnelle sous sa forme finale :

$$\vec{y} = \vec{x} - \vec{\alpha}(\vec{x}). \quad (1.6.8)$$

Conformément à l'Équation 1.6.8, la connaissance de la position intrinsèque  $\vec{y}$  d'un objet dans le plan source ne permet pas de déterminer directement sa position apparente  $\vec{x}$ , car  $x$  ne peut être isolé. C'est pourquoi, dans le contexte des lentilles gravitationnelles, nous faisons appel à une méthode de rétro-propagation. Cette méthode initie le processus par la position apparente pour déduire la position intrinsèque. En pratique, pour simuler les effets d'une lentille gravitationnelle, nous divisons le ciel observé en pixels et pour chaque pixel, nous résolvons l'équation de lentille pour déterminer d'où il provient. Nous utilisons ensuite cette cartographie pour déterminer où les images d'une source devraient apparaître. Cette méthode de résolution se nomme la simulation par traçage des rayons (de l'anglais *ray-tracing*

*simulation*). Puisque l'équation de lentille est hautement non-linéaire, une source peut mener à une multiplicité d'images apparentes.

### 1.6.2. Potentielle de la lentille

Pour résoudre un système de lentille, il faut un modèle physique qui modélise simultanément la position intrinsèque de la source et la distribution de masse de la lentille. Cette dernière est caractérisée par le potentiel gravitationnel effectif de la lentille. Celui-ci est obtenu en projetant son potentiel newtonien tridimensionnel,  $\phi(D_L\vec{\theta}, z)$ , sur le plan lentille. Pour ce faire, il suffit d'intégrer  $\phi(D_L\vec{\theta}, z)$  le long de l'axe  $z$ , qui est celui de la ligne de visée. En rajoutant un facteur d'échelle, on obtient le potentiel effectif,  $\hat{\Psi}(\vec{\theta})$ , suivant :

$$\hat{\Psi}(\vec{\theta}) = \frac{2D_{LS}}{D_L D_S c^2} \int \phi(\vec{\theta}, z) dz. \quad (1.6.9)$$

Celui-ci satisfait à deux propriétés fondamentales. D'abord, son gradient est égal à l'angle de déviation :

$$\vec{\nabla}_{\theta} \vec{\Psi}(\vec{\theta}) = \vec{\alpha}(\vec{\theta}). \quad (1.6.10)$$

Secondement, son Laplacien équivaut à deux fois la convergence :

$$\Delta_{\theta} \vec{\Psi}(\vec{\theta}) = 2\kappa(\vec{\theta}). \quad (1.6.11)$$

La convergence,  $\kappa(\vec{\theta})$ , est défini comme la surface de densité adimensionnelle de la lentille,

$$\kappa(\vec{\theta}) \equiv \frac{\Sigma(\vec{\theta})}{\Sigma_c(\vec{\theta})}, \quad (1.6.12)$$

$$= \frac{4\pi G}{c^2} \frac{D_L D_{LS}}{D_S} \Sigma(\vec{\theta}), \quad (1.6.13)$$

où  $\Sigma_c(\vec{\theta})$  représente la densité critique, une quantité caractéristique du système de lentille qui dépend des distances angulaires.  $\Sigma(\vec{\theta})$  décrit la surface de densité, une fonction qui découle de l'intégrale en  $z$  du Laplacien du potentiel newtonien tridimensionnel :

$$\Sigma(\vec{\theta}) = \frac{1}{4\pi G} \int \Delta\phi(\vec{\theta}, z) dz. \quad (1.6.14)$$

À noter que cette deuxième propriété (Équation 1.6.11) découle initialement de l'équation de Poisson :  $\Delta\phi = 4\pi G\rho$ .

### 1.6.3. Grossissement et étirement local

L'effet d'une lentille gravitationnelle forte induit une distorsion non linéaire de l'objet observé. Cependant, lorsque la variation de position dans le plan de la lentille est suffisamment limitée, cette déformation peut être approximée de manière linéaire. On parle alors d'un effet de lentille localement linéaire. Cette approche de résolution est qualifiée de premier ordre, car elle implique la résolution d'une équation différentielle linéaire de premier ordre.

Considérons un point sur le plan lentille à  $\vec{\theta}_0$  où l'angle de déviation est de  $\vec{\alpha}_0$ . Localement, on trouve que

$$\vec{\theta} = \vec{\theta}_0 + d\vec{\theta}, \quad (1.6.15)$$

$$\vec{\beta} = \vec{\beta}_0 + d\vec{\beta}, \quad (1.6.16)$$

$$\vec{\alpha} = \vec{\alpha}_0 + \frac{d\vec{\alpha}}{d\vec{\theta}} d\vec{\theta}, \quad (1.6.17)$$

en remplaçant ces équations dans l'équation de lentille, on obtient

$$(\vec{\beta} - \vec{\beta}_0) = \left( I - \frac{d\alpha(\vec{\theta})}{d\vec{\theta}} \right) (\vec{\theta} - \vec{\theta}_0). \quad (1.6.18)$$

Ainsi, ayant défini l'effet de la lentille gravitationnelle comme étant localement linéaire,  $\vec{y} = A\vec{x}$ , la distorsion de l'image peut être exprimée par l'intermédiaire de la matrice Jacobienne suivante :

$$A \equiv \frac{\partial \vec{\beta}}{\partial \vec{\theta}} = \left( \delta_{ij} - \frac{\partial \alpha_i(\vec{\theta})}{\partial \theta_j} \right) = \left( \delta_{ij} - \frac{\partial^2 \Psi(\vec{\theta})}{\partial \theta_j \partial \theta_i} \right). \quad (1.6.19)$$

Ici,  $i$  et  $j$  symbolisent les composantes des vecteurs de position dans le plan de la lentille, se limitant à deux dimensions.

Avec la formule de la matrice de distorsion à notre disposition, nous sommes en mesure de caractériser les diverses formes de distorsion générées par les lentilles gravitationnelles. L'étirement est une déformation primordiale induite par l'effet de lentille. Il est responsable de la création de larges arcs lumineux en raison de son anisotropie : un axe est dilaté tandis que l'autre est affiné. Dans la matrice de distorsion,  $A$ , l'étirement est gouverné par les éléments non diagonaux :

$$\left( A - \frac{1}{2} \text{tr} A \cdot I \right)_{ij} = \delta_{ij} - \frac{\partial^2 \Psi(\vec{\theta})}{\partial \theta_j \partial \theta_i} - \frac{1}{2} (1 - \Psi_{11} + 1 - \Psi_{22}) \delta_{ij}, \quad (1.6.20)$$

$$= -\frac{\partial^2 \Psi(\vec{\theta})}{\partial \theta_j \partial \theta_i} + \frac{1}{2} (\Psi_{11} + \Psi_{22}) \delta_{ij}. \quad (1.6.21)$$

Les éléments diagonaux de la matrice, quant à eux, sont responsables de la magnification, une déformation isotrope. Le facteur de magnification, noté  $\mu$ , peut être calculé directement à partir de la matrice de distorsion locale comme suit :

$$\mu = \frac{1}{\det A}. \quad (1.6.22)$$

Le grossissement, noté  $\mu$ , permet de quantifier les modifications de la taille apparente et de la luminosité d'une source induites par l'effet de lentille. Si  $|\mu| > 1$ , l'image est amplifiée ; tandis qu'elle est réduite si  $0 < |\mu| < 1$ . Il est à noter que le signe de  $\mu$  porte une signification physique : un  $\mu$  positif indique que l'image n'est pas inversée par rapport à la source, alors qu'un  $\mu$  négatif signifie une inversion de l'image. Il est essentiel de rappeler que cette

approche de résolution linéaire est seulement valide à une échelle locale. Dans le contexte des lentilles gravitationnelles fortes, où de multiples images sont générées, cette méthode doit être appliquée individuellement pour chaque image. Ainsi, il est possible d’exploiter chaque image, chacune étant déformée différemment.

Les distorsions induites par les effets des lentilles gravitationnelles fortes offrent une opportunité d’analyser des objets extrêmement éloignés. Dans certains cas, ces objets sont si distants que leur émission lumineuse se situe en dessous du seuil de détection du télescope. Toutefois, la magnification et l’étirement causés par la lentille gravitationnelle peuvent amplifier ces émissions jusqu’à un niveau détectable (Bayliss et al. 2020). Dans le contexte de ce projet, une galaxie contenant un SMBH en son centre est observée à différentes fréquences puis distordue par une masse en avant-plan. Étant donné que l’effet de lentille est achromatique, la reconstruction de l’image à chaque fréquence peut être effectuée de manière indépendante et parallèle. Le produit final est un cube de données contenant l’image distordue des lignes d’émission du gaz moléculaire situé dans la galaxie et autour du SMBH. Cette distorsion est exploitée pour optimiser la résolution spatiale de ces émissions, qui constitue le facteur limitant dans la mesure directe de la masse des SMBHs (Davis et al. 2017; Smith et al. 2019; Lelli et al. 2022; Ruffa et al. 2023). Il devient ainsi possible de mesurer la masse des SMBHs à des distances sans précédent.

## 1.7. L’apprentissage automatique

L’apprentissage automatique est une branche de l’intelligence artificielle (AI, de l’anglais *Artificial Intelligence*) qui permet le développement d’algorithmes capable d’apprendre à partir de données et de prendre des décisions sans avoir besoin d’être explicitement programmés. L’AI est devenu de plus en plus populaire, car elle permet d’automatiser des tâches complexes et de résoudre des problèmes pour lesquels les approches traditionnelles ne sont pas adaptées. Les données jouent un rôle crucial dans l’apprentissage automatique, et les algorithmes doivent être entraînés sur de grandes quantités de données de haute qualité pour produire des résultats non biaisés. Dans le domaine de l’astronomie, l’arrivée de nouveaux télescopes tels que LSST (Ivezić et al. 2019) et Euclide (Euclid Collaboration et al. 2022) souligne l’importance de développer des outils d’apprentissage automatique pour étudier la quantité phénoménale de données brutes qui seront récoltées.

### 1.7.1. Les réseaux de neurones artificiels

Les réseaux de neurones, également désignés sous le nom de réseaux neuronaux artificiels (ANN, pour *Artificial Neural Network* en anglais), constituent une branche de l’intelligence artificielle dont le fonctionnement est inspiré par la structure et les mécanismes des neurones



biologiques (Goodfellow et al. 2013). Un réseau neuronal se compose d'une couche d'entrée, d'un nombre prédéterminé de couches cachées et d'une couche de sortie. Les données sont initialement introduites par la couche d'entrée, dont le nombre de neurones correspond au nombre de paramètres initiaux. Les couches cachées modifient ensuite l'information en fonction de paramètres ajustables. Finalement, la couche de sortie délivre la valeur finale produite par le réseau.

Considérons un vecteur de données initiales  $\vec{s}$  de dimension  $S$  qui est transmis à une couche cachée composée de  $H$  neurones. Chaque neurone de cette première couche cachée, noté  $h_{ij}$ , reçoit l'information sous la forme d'une combinaison linéaire des données initiales :

$$h_j^1 = \sum_{i=1}^S w_{ji}^1 s_i + b_j^1. \quad (1.7.1)$$

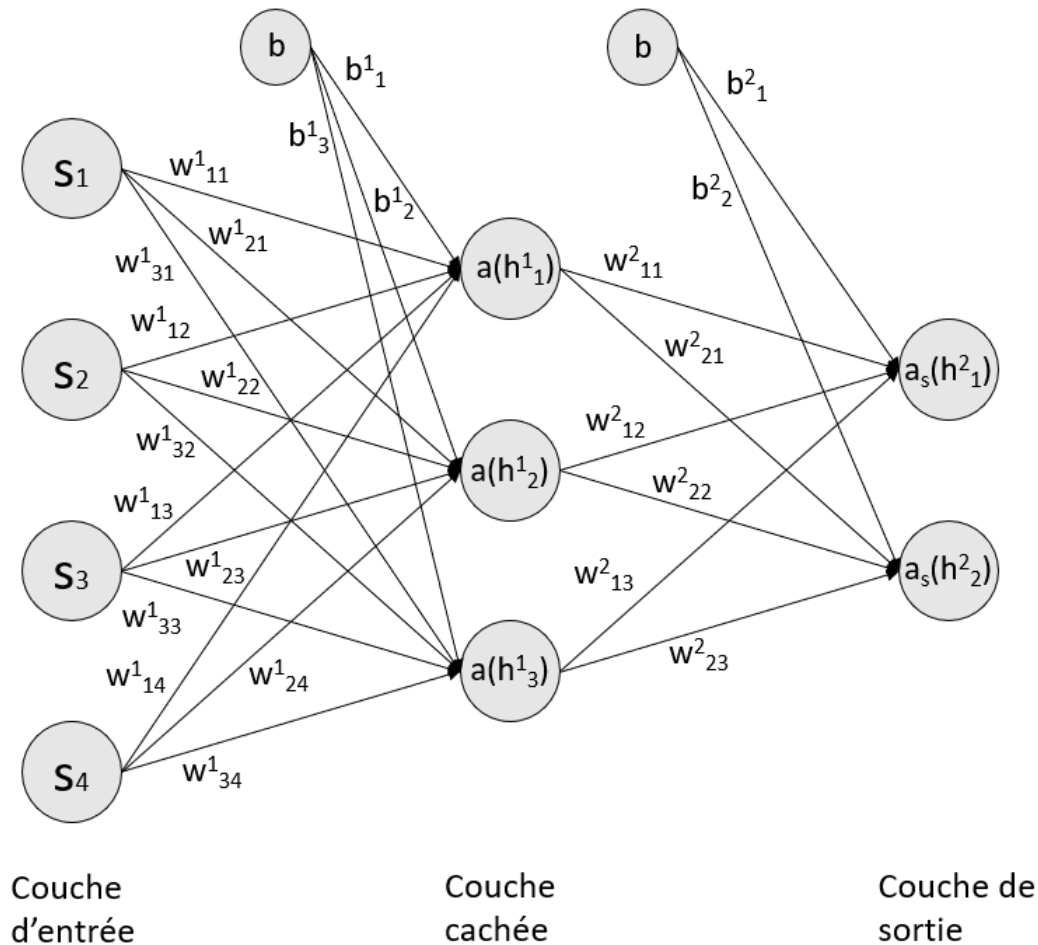
Dans cette expression,  $w_{ji}^1$  représente les poids associés à chaque connexion, tandis que  $b_j^1$  fait référence au biais appliqué à chaque neurone. Ici, l'indice  $i \in [1, S]$  désigne une dimension d'entrée spécifique et  $j \in [1, H]$  identifie un neurone particulier de la couche cachée. Quant à l'exposant, il précise à quelle couche cachée la notation se réfère.

Pour injecter de la non-linéarité dans le processus de propagation de l'information à travers le réseau, une fonction d'activation  $a$  est appliquée à chaque neurone (Clevert et al. 2015). Cette fonction doit être à la fois non-linéaire et différentiable. Plusieurs types de fonctions d'activation sont couramment utilisés, notamment la sigmoïde, ReLU et ELU. Il est essentiel de préciser que le choix de la fonction d'activation pour la couche de sortie dépend de la nature du problème à résoudre. La fonction sigmoïde s'adapte bien à la prédiction de probabilité puisque les valeurs de l'image sont comprise entre 0 et 1 alors que les fonctions ReLU et ELU sont plutôt favorisé pour la régression. En combinant les poids, les biais et les fonctions d'activation, avec un nombre suffisant de neurones et de couches cachées, il devient théoriquement possible d'approcher n'importe quelle fonction lisse avec une précision arbitraire à l'aide du réseau (Bengio et al. 2015).

La Figure 1.10 illustre un réseau de neurones comportant une unique couche cachée. Le vecteur de sortie  $\vec{y}_{prédit}$  présente une dimension  $Z = 2$ . La fonction suivante représente le réseau dans son intégralité :

$$y_{prédit_k} = a_s \left( \sum_{j=1}^H w_{kj}^2 a \left( \sum_{i=1}^S w_{ji}^1 s_i + b_j^1 \right) + b_k^2 \right), \quad (1.7.2)$$

où  $k = [1, Z]$  et  $a_s$  désigne la fonction d'activation de la couche de sortie. Comme mentionné précédemment, cette fonction peut différer de  $a$  (la fonction d'activation des neurones cachés), en fonction de la nature du problème à résoudre.



**Figure 1.10** – Réseau de neurones pour un problème de régression comportant 5 paramètres initiaux et une valeur de sortie. L'information se déplace donc de gauche à droite. Le choix du nombre de couches internes et du nombre de neurones dans chaque couche est complètement arbitraire. Il s'agit d'hyperparamètres qui doivent être optimisés en comparant la performance du réseau pour différentes valeurs.

Une fois le réseau établi, un objectif est nécessaire. En termes mathématiques, cet objectif se manifeste sous la forme d'une fonction de perte,  $L$ . Le réseau sera formé pour minimiser cette perte autant que possible. Dans le cadre d'un apprentissage supervisé, le réseau peut se former en utilisant des exemples de paires d'entrées et de sorties. La tâche du réseau est donc de prédire la sortie correcte (Mathieu et al. 2015). L'une des fonctions de perte les plus couramment utilisées dans ce type d'apprentissage est l'erreur quadratique moyenne (MSE – de l'anglais *Mean Squared Error*) :

$$L(y_{\text{prédit}}, y_{\text{étiquette}}) = \frac{1}{N} \sum_{n=1}^N (y_{\text{prédit}_n} - y_{\text{étiquette}_n})^2, \quad (1.7.3)$$

où  $N$  représente le nombre total de données d'entrée,  $y_{\text{prédit}}$  correspond à la valeur prédite par le réseau et  $y_{\text{étiquette}}$  à la réponse correcte. Trouver les valeurs de  $w$  et  $b$  qui minimisent le résultat de cette équation a pour effet d'améliorer la valeur prédite par le réseau. Néanmoins, cette fonction de perte requiert des données étiquetées, c'est-à-dire que nous devons connaître la réponse correcte à l'avance afin de former le réseau à prédire le résultat adéquat.

Il existe plusieurs méthodes d'optimisation visant à minimiser la fonction de perte. Parmi ces méthodes, on peut citer la descente stochastique du gradient (SGD - de l'anglais *Stochastic Gradient Descent*), Adam et Adamax qui s'appuient toutes sur le principe de rétropropagation du gradient. En calculant le gradient depuis la fonction de perte jusqu'aux données d'entrée, il est possible d'identifier les ajustements numériques à apporter aux poids et aux biais pour minimiser davantage le coût. Ces ajustements sont guidés par le taux d'apprentissage : plus celui-ci est élevé, plus grand sera le pas dans la direction indiquée par le gradient et vice-versa. Cette étape est répétée plusieurs fois jusqu'à convergence.

Toutefois, le seul objectif de la descente du gradient étant de minimiser la perte, le réseau risque de ne pas apprendre à généraliser le problème, mais plutôt de mémoriser les paires de données d'entrée et de sortie (Goodfellow et al. 2016). Pour garantir que le réseau apprend des caractéristiques pertinentes applicables à des données inédites, l'ensemble de données initial est divisé en trois : un ensemble d'entraînement, un ensemble de validation et un ensemble de test. L'apprentissage est effectué sur l'ensemble d'entraînement qui constitue la majorité des données. Parallèlement à l'entraînement, la performance du réseau est évaluée sur l'ensemble de validation. Cette démarche permet de vérifier que le réseau fonctionne bien sur des données qui n'ont pas été utilisées pour optimiser la fonction de perte. De plus, l'ensemble de validation guide aussi les hyperparamètres, tels que la taille du réseau (par exemple, le nombre de couches cachées) et le taux d'apprentissage (LeCun et al. 2015). Enfin, pour connaître la performance réelle du réseau, celui-ci doit être testé sur des données qui n'ont jamais été vues ni utilisées de quelque manière que ce soit pour améliorer la performance du réseau. C'est là qu'intervient l'ensemble de test.

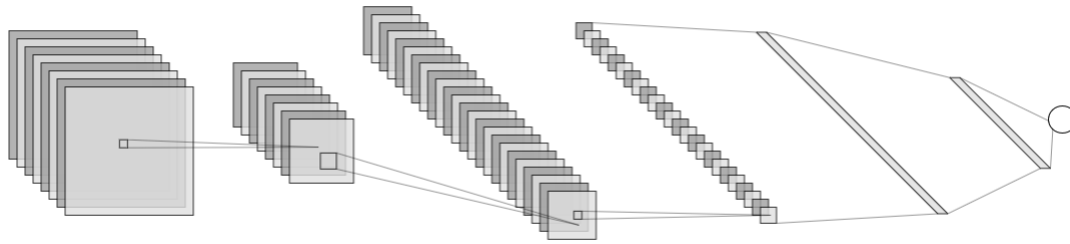
Une fois entraîné, le réseau de neurones peut être utilisé pour prendre des décisions ou effectuer des prédictions en temps réel. Contrairement aux méthodes conventionnelles, qui nécessitent du temps pour converger vers la prédiction correcte pour chaque nouvelle observation, un réseau de neurones entraîné peut réduire le temps d'analyse des données de plusieurs ordres de grandeur.

### 1.7.2. Les réseaux de neurones convolutifs

Les réseaux de neurones convolutionnels (CNN – de l’anglais *Convolutional Neural Network*) sont particulièrement adaptés aux tâches de traitement d’images, grâce à leur architecture spécifique qui exploite l’aspect spatial de ces données (Simonyan & Zisserman 2014). Les CNNs mettent en œuvre une opération distinctive nommée *convolution*. Cette dernière correspond à un filtre qui parcourt l’image en multipliant la valeur des pixels par celles correspondantes au filtre, suivies de leur sommation. Ce processus contribue à l’identification de caractéristiques ou motifs dans les images, tels que les bords, les textures ou les couleurs (Szegedy et al. 2014). Les filtres sont déterminés au cours de l’entraînement, permettant au réseau d’extraire les caractéristiques les plus pertinentes pour l’objectif visé.

Les CNNs démontrent une propriété d’invariance spatiale, ce qui signifie qu’ils peuvent identifier une caractéristique, quel que soit son emplacement dans l’image. Pour accomplir cela, le réseau utilise un ensemble de filtres qui se déplacent de manière systématique à travers l’image. Chaque filtre, qui est essentiellement une petite matrice de valeurs, effectue une opération de convolution lors de son déplacement sur l’image. Cette opération consiste en une multiplication élément par élément avec les pixels de l’image qu’il couvre, suivie de la sommation de ces produits pour obtenir une valeur unique. Cette procédure génère une nouvelle image, dénommée *carte de caractéristiques*, qui est ensuite transmise à la couche suivante du réseau. Ce processus, qui englobe les étapes de convolution suivies ultérieurement par un sous-échantillonnage (destiné à réduire la dimension spatiale), permet au réseau de reconnaître des motifs, indépendamment de leur localisation ou de leur taille dans l’image. Les images constituent des données de haute dimensionnalité. Par exemple, une image couleur de 100x100 pixels comprend 30,000 valeurs de pixels à traiter. Les CNNs simplifient la complexité des images par le biais des opérations de convolution et de sous-échantillonnage, ce qui a pour effet de diminuer le nombre de paramètres à apprendre en comparaison avec un réseau de neurones entièrement connecté (He et al. 2015). Par conséquent, le processus d’apprentissage est optimisé.

Les CNNs sont largement exploités en astrophysique, compte tenu de leur aptitude à traiter des données présentant une invariance par translation. En effet, des CNNs ont été employés pour la classification de différents types de galaxies (Dieleman et al. 2015; Huertas-Company et al. 2015; Tadaki et al. 2020; Cheng et al. 2021; Villanueva-Domingo et al. 2022; Remy et al. 2023) ainsi que pour l’inférence des propriétés physiques galactiques (Hoyle 2016; Tuccillo et al. 2018; Pasquet et al. 2019; Eriksen et al. 2020; Li et al. 2022). Ils ont aussi servi à modéliser les effets des lentilles gravitationnelles fortes (Perreault Levasseur et al. 2017; Hezaveh et al. 2017b; Morningstar et al. 2019; Pearson et al. 2019; Bom et al. 2019; Madireddy et al. 2019; Charnock et al. 2020; Hausen & Robertson 2022), parmi d’autres



**Figure 1.11** – Représentation schématique de l’architecture d’un CNN, un type de réseau de neurones spécialement adapté pour le traitement des images. Chaque couche du CNN analyse une partie de l’image, extrait des caractéristiques et réduit la résolution spatiale de celle-ci. Une fois les caractéristiques spatiales encodées, l’image est écrasée puis passée dans un réseau neuronal classique pour produire une réponse finale.

applications variées. Dans le contexte de notre projet, nous utiliserons un CNN afin de traiter des cubes de données ALMA et estimer la masse de SMBHs.

### 1.7.3. L’augmentation des données

L’augmentation de données est une technique courante dans le traitement d’images qui consiste à générer de nouvelles données d’entraînement en appliquant des transformations diverses sur l’ensemble de données existant. Cette méthode vise à accroître la diversité et la quantité des données, ce qui peut améliorer les performances des modèles d’apprentissage automatique (Bengio et al. 2013). Les transformations d’images courantes incluent le retournement, la rotation, le redimensionnement, le recadrage et l’ajout de bruit aux images. Ces transformations aident le modèle à mieux généraliser et apprendre les symétries du problème en le soumettant à une plus grande variété de variations dans les données d’entrée. L’augmentation de données est un outil puissant qui est devenu une pratique courante dans la classification d’images, la détection d’objets et d’autres tâches de vision par ordinateur.

### 1.7.4. Les incertitudes

Une contrainte couramment rencontrée lors de l’emploi de l’apprentissage automatique concerne l’absence d’incertitude sur les prédictions, particulièrement dans le cadre des problèmes de régression (Nair et al. 2022). Une solution envisageable à cette problématique consiste à utiliser le négatif du logarithme naturel de l’équation gaussienne comme fonction de perte (Équation 1.7.4).

$$L(y_{\text{prédit}}, y_{\text{étiquette}}, \sigma) = -\frac{1}{N} \sum_{n=1}^N \ln \left[ \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left( -\frac{(y_{\text{prédit}_n} - y_{\text{étiquette}_n})^2}{2\sigma^2} \right) \right], \quad (1.7.4)$$

$$= \frac{1}{N} \sum_{n=1}^N \left( \frac{(y_{\text{prédit}_n} - y_{\text{étiquette}_n})^2}{2\sigma^2} + \pi\sigma^2 \right). \quad (1.7.5)$$

La fonction de perte gaussienne est largement utilisée en apprentissage automatique pour les problèmes de prédiction de variables de sortie continues (Mlotshwa et al. 2023). En plus de minimiser l'écart entre les valeurs prédites et les valeurs réelles, cette fonction de perte permet également au modèle de prédire l'incertitude associée à ses prévisions. Cela est réalisé en modélisant la sortie sous forme d'une distribution gaussienne caractérisée par une moyenne ( $y_{\text{prédit}}$ ) et un écart type ( $\sigma$ ). La moyenne représente l'estimation ponctuelle, tandis que l'écart type représente l'incertitude de la prédiction. En utilisant la fonction de perte gaussienne, le modèle peut apprendre à fournir des estimations d'incertitude plus élevées dans les zones de l'espace d'entrée où il a moins de données ou où les données sont plus ambiguës. Toutefois, il convient de noter que cette méthode approxime la distribution de la variable prédite comme étant gaussienne, ce qui n'est pas toujours le cas.

## 1.8. Description de ce mémoire

Cette section a pour but de déclarer explicitement et avec plus de détail les objectifs et méthodes employées au cours de ce mémoire.

### 1.8.1. Objectifs

L'objectif primaire de ce mémoire est de développer un outils basé sur l'AI capable de tirer profit de l'effet des lentilles gravitationnelles pour mesurer la masse de SMBHs à des distances jamais fait auparavant. Dans cette optique, nous générons d'abord des cubes de données ALMA issue de la simulation FIRE pour entraîner un CNN à automatiquement extraire la masse du SMBH. Une fois assez de données générés et le CNN entraîné, on apprend le score de celles-ci, pour ensuite pouvoir, par inférence bayésienne, reconstruire le cube distordu par l'effet de lentille gravitationnelle. Il est alors possible d'appliquer directement le CNN préalablement entraîné pour mesurer la masse des cubes reconstruits. Cette méthode permet non seulement de tirer profit de l'effet de grossissement des lentilles gravitationnelles pour mesurer la masse de SMBHs à des distances plus élevées, mais est aussi totalement automatique.

## 1.9. Déclaration de l'étudiant

Je, David Chemaly, déclare, à moins d'indication contraire, que la totalité de ce mémoire ainsi que le travail présenté ont été effectués par moi-même. La rédaction complète du mémoire, tous chapitres confondus, a été réalisée par l'auteur. Dans les courts paragraphes qui suivent, je décris plus amplement la contribution des autres auteurs.

L'idée de ce projet a été initialement proposée par Hezaveh 2014, qui a étudié la possibilité de résoudre la SOI de SMBHs à de plus grandes distances en utilisant la distorsion des lentilles gravitationnelles. Sous la co-supervision de Yashar Hezaveh, expert en IA, et Julie Hlavacek-Larrondo, experte en SMBH, j'ai mené à bien ce projet.

Les données utilisées pour entraîner le CNN ont été fournies par la suite de simulations hydrodynamiques FIRE. La collaboration FIRE est dirigée par Phil Hopkins, Claude-André Faucher-Giguère, Dusan Keres et Eliot Quataert, qui nous ont donné accès à des simulations privées. Daniel Anglés-Alcázar a facilité la communication entre l'étudiant et le reste de la collaboration, ainsi que la mise à disposition de tous les éléments nécessaires à l'utilisation des simulations.

Pour convertir les simulations en cubes ALMA réalistes, Axel Weiss a partagé avec l'étudiant un logiciel permettant de passer des paramètres de la simulation (ex. : vitesse, température, pression, etc.) à une valeur de flux observée.

Alexandre Adam a été responsable de l'apprentissage du score de distribution des données et de la reconstruction des cubes ALMA distordus par l'effet de lentilles gravitationnelles.

Mario Pasquato a apporté son aide et ses commentaires de manière régulière tout au long du projet. Vassily Belokurov et Roberto Maiolino ont également fait de même pendant un stage de trois mois suivi par l'étudiant à l'Université de Cambridge.

## Chapitre 2

---

# Measuring the mass of supermassive black holes at higher redshifts through gravitational lensing with machine learning

David Chemaly,<sup>1,2,3</sup> Julie Hlavacek-Larrondo,<sup>1,2,3</sup> Yashar Hezaveh,<sup>1,2,3,5</sup> Alexandre Adam,<sup>1,2,3</sup> Mario Pasquato,<sup>1,3,4,5</sup> Axel Weiß<sup>6</sup> and Daniel Anglés-Alcázar<sup>7,8</sup>

<sup>1</sup>*Département de physique, Université de Montréal, C.P. 6128 Succ. Centre-ville, Montréal H3C 3J7, Canada*

<sup>2</sup>*Centre de recherche en astrophysique du Québec (CRAQ)*

<sup>3</sup>*Ciela, Computation and Astrophysical Data Analysis Institute, Montreal, Quebec, Canada*

<sup>4</sup>*Physics and Astronomy Department Galileo Galilei, University of Padova, Vicolo dell'Osservatorio 3, I-35122, Padova*

<sup>5</sup>*Mila - Quebec Artificial Intelligence Institute, Montreal, Quebec, Canada*

<sup>6</sup>*Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69 D-53121 Bonn, Germany*

<sup>7</sup>*Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA*

<sup>8</sup>*Department of Physics, University of Connecticut, 196 Auditorium Road, U-3046, Storrs, CT 06269-3046, USA*

Cet article est en cours de préparation pour soumission à la revue *Monthly Notices of the Royal Astronomical Society* (MNRAS).



# Abstract

Supermassive black holes (SMBHs) are integral to the coevolution of their host galaxies and significantly influence their physical properties. However, accurately determining the masses of SMBHs remains a substantial challenge, even within the local universe. Currently, the most dependable method involves examining the kinematics of molecular gas within the sphere of influence (SOI) of an SMBH, where the gravitational potential of the black hole overwhelms that of the host galaxy. Observing this relatively compact region, which usually spans a mere few tens of parsecs, requires high spatial resolution, thereby posing a significant obstacle. Consequently, employing this approach to gauge SMBH masses at high redshifts is presently considered unattainable for the foreseeable future. To examine the feasibility of harnessing machine learning to leverage the magnification from strong gravitational lensing for SMBH mass measurements at higher redshifts, it is essential to first generate a reliable and sufficient volume of data. In this study, we introduce an innovative fast and automated pipeline to simulate observations of CO(2-1) molecular gas flux as would be captured by the Atacama Large Millimeter/submillimeter Array (ALMA). The preliminary mock ALMA data is generated from a zoomed-in cosmological simulation, *Feedback in Realistic Environments* (FIRE). The kinematic effects of the SMBH are incorporated by assuming a virialized molecular cloud. The consequential emitted flux is calculated as a function of the gas’s velocity, temperature, density, H<sub>2</sub> fraction, redshift, and its apparent size in the sky. Strong gravitational lensing effects are modeled through ray-tracing, with the assumption of a singular isothermal ellipsoid (SIE) mass profile. The ALMA data cube is then generated, factoring in the spatial and spectral resolution, determined by the number of antennas, their configuration, and the exposure time. We subjected the generative pipeline to rigorous testing on the task of predicting SMBH mass from unlensed data. Consistent with conventional methods and corroborated by saliency maps, our network utilizes information within the SOI to make its mass predictions. If the spatial resolution fails to adequately resolve the SOI, then it is found that the predictions are inaccurate.

**Keywords :** quasars : supermassive black holes – gravitational lensing : strong – black hole physics – galaxy : kinematics and dynamics

## 2.1. Introduction

The scientific community has established a well-documented correlation between the mass of supermassive black holes (SMBHs;  $M_{BH}$ ) and various galaxy-scale parameters such as luminosity and mass (Magorrian et al. 1998; Kormendy & Ho 2013a; Graham 2016). Among these correlations, the  $M_{BH}-\sigma_*$  relation (Ferrarese & Merritt 2000a; Gebhardt et al. 2000a), where  $\sigma_*$  denotes the bulge velocity dispersion, stands out as the most robust with the less

scatter, indicating a strong co-evolution between galaxies and the formation of black holes throughout the history of our universe. Understanding the underlying physics driving this relationship remains a fascinating challenge for astrophysicists today.

Over the years, various direct methods have been employed to obtain accurate measurements of SMBH masses, including kinematical tracers based on masers (Greene et al. 2010; Kuo et al. 2011), on stars (Ghez et al. 2008; Walsh et al. 2013), ionized gas (Ford et al. 1994; Marconi et al. 2006), or molecular gas (Riffel & Storchi-Bergmann 2011; Davis et al. 2013) motions and reverberation mapping (Kaspi et al. 2000; Peterson et al. 2004). Among these tracers, stellar motion measurements tend to introduce more uncertainties compared to gas measurements due to their point-like nature. Within the gas category, molecular gas, with its lower temperature, tends to exhibit less chaotic behavior than ionized gas, enabling more precise measurements. However, when utilizing kinematic measurements to reliably determine the mass of a SMBH, it is crucial to spatially resolve the sphere of influence (SOI) of the SMBH. The SOI is defined as the region where the gravitational potential of the SMBH dominates over that of the galaxy. The radius of this sphere can be approximated as follows :

$$r_{SOI} \approx \frac{GM_{BH}}{\sigma_{\star}}, \quad (2.1.1)$$

where  $G$  is Newton's gravitational constant. Depending on the black hole mass, the size of the SOI can range from a few tens of parsecs to a few hundreds. Consequently, to measure the mass of a SMBH using kinematics, observational data with resolution down to parsec scales is necessary.

With the advancements in angular resolution and sensitivity of (sub-)millimeter interferometry, multiple studies have been successful in measuring the mass of SMBHs by studying the dynamical rotation of cold molecular gas through CO emission lines (e.g. Onishi et al. 2017; Barth et al. 2016a; Smith et al. 2019). The mm-Wave Interferometric Survey of Dark Object Masses (WISDOM) project has been at the forefront of leveraging high-resolution CO observations from the Atacama Large Millimeter/submillimeter Array (ALMA) to achieve this goal. To date, the WISDOM project and other research teams have utilized this method to measure the masses of 12 early-type galaxies (e.g. Davis et al. 2017, 2018b; North et al. 2019; Boizelle et al. 2019; Smith et al. 2021b), 7 late-type galaxies with AGN (Combes et al. 2019), as well as a barred spiral galaxy (Nguyen et al. 2020). These studies have consistently demonstrated the precision, accuracy, and robustness of this method in measuring the masses of different types of SMBHs in various galaxies.

However, the primary limitation lies in the required spatial resolution to probe the dynamical rotation. Considering the best angular resolution of ALMA when observing CO(2-1) emission lines, the sphere of influence of an extremely massive black hole of  $10^{10}M_{\odot}$  could only be resolved at a redshift no greater than 0.2. Thus, given the current generation of telescopes, measuring the masses of SMBHs through molecular gas kinematics at greater redshifts requires some form of spatial magnification.

Strong gravitational lensing acts as a natural telescope, focusing the emitted flux of a background object and magnifying its angular size. This achromatic phenomenon affects photons of all wavelengths equally, making it a valuable tool in various fields of astronomy to study objects that would otherwise be inaccessible. Notably, it has been employed to identify the oldest galaxies in the infrared (Atek et al. 2023), quantify the star formation rate of young galaxies at high redshift through X-ray emission (Bayliss et al. 2020), and explore other important areas of research (e.g. Hoekstra et al. 2004; Hezaveh et al. 2016). Thus, strong gravitational lensing can be harnessed to enhance the angular resolution of ALMA observations, enabling spatial resolution of the SOI at high redshifts.

It is crucial to acknowledge that the advantages of strong gravitational lensing come with a trade-off : the observed image is significantly distorted. This non-linear transformation can be reversed through conventional sampling methods, such as Monte Carlo Markov Chain (MCMC) techniques (Foreman-Mackey et al. 2013), but it requires making simplifying assumptions (e.g., stable and symmetric circular orbit) and extensive computational resources.

In recent years, the remarkable advancements in graphics processing unit (GPU) computation power, coupled with the vast amounts of data that need to be processed, have propelled machine learning to the forefront of astronomical research. Deep neural networks have exhibited remarkable capabilities in assisting astronomers in unraveling longstanding mysteries of our universe. They have been instrumental in identifying the primary driver of galaxy quenching (Piotrowska et al. 2022), providing an independent means of measuring the Hubble constant (Huber et al. 2022), and rapidly classifying different types of galaxies from survey data (Domínguez Sánchez et al. 2018). In the context of strong gravitational lensing, recent work, such as Adam et al. 2022, have demonstrated the effectiveness of score based models in reconstructing the posterior of non-linearly distorted background images. In this study, we build upon the work of Hezaveh 2014 to generate a fast and automated pipeline to create realistic mock ALMA data of the CO(2-1) emission line from the *Feedback In Realistic Environment-2* (FIRE, Hopkins et al. 2014a). The pipeline is then test on on the regressive task of predicting the posterior distribution of the SMBH mass from the simulated data.

The layout of this paper is as follows : Section 2 introduces the data utilized in this study. Section 3 details the processes involved in acquiring highly realistic, simulated ALMA

data of galaxies hosting a SMBH. The Convolutional Neural Network (CNN) employed for inferring the SMBH mass is delineated in Section 4. Section 5 presents our discussion, which is then followed by a conclusion in Section 6. This final section encapsulates the key findings of our research and underscores potential avenues for future exploration.

## 2.2. Data

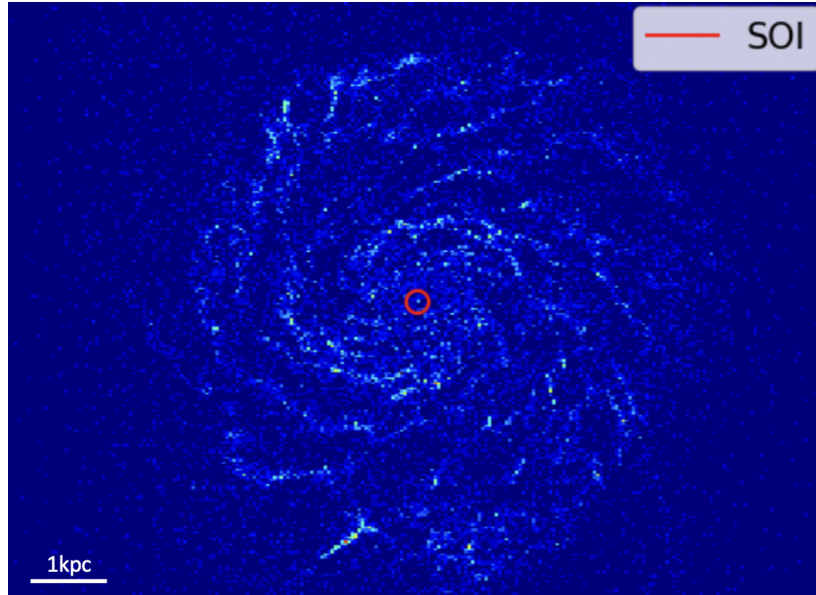
### 2.2.1. FIRE simulations

The end goal of this study is to train a machine learning network capable of accurately measuring the mass of SMBHs using ALMA observations. To accomplish this goal, we will employ a supervised learning approach that necessitates prior knowledge of the SMBH masses. However, due to the limited availability of real SMBH mass measurements obtained from molecular gas observations with ALMA, it is necessary to develop a customized pipeline for generating mock ALMA observations based on simulations.

We utilize simulations from the FIRE-2 data release (Wetzel et al. 2023). These cosmological zoom-in simulations, part of the FIRE project, were generated using the Gizmo code (Hopkins 2015) and the FIRE-2 physics model (Hopkins et al. 2018). The FIRE simulations incorporate comprehensive treatments of various important physical processes, including gas cooling, star formation, stellar feedback, and the influence of stellar radiation and supernova explosions. By explicitly accounting for these processes, the FIRE simulation aims to accurately reproduce the observed properties of galaxies across a wide range of masses and redshifts. In this work, we provide a brief overview of the simulation methodology and refer to Hopkins et al. 2018 for a more complete description.

All the FIRE-2 simulations adopt a zoom-in technique, focusing on specific regions embedded within a lower-resolution cosmological framework (Hopkins et al. 2014a). This approach allows for a high spatial resolution that is crucial for our study as we aim to resolve the SOI of SMBHs. By zooming in on selected regions, the simulations provide the necessary detail to capture the dynamics and physical properties of the SMBHs and their surrounding environments at a scale suitable for our analysis (see Figure 2.1).

In this project, our focus is primarily on the *MassiveFIRE* catalogue within the FIRE-2 simulations due to its inclusion of models for the growth of massive black holes through gravitational torques between the gas and stellar components (Anglés-Alcázar et al. 2017a). However, it is important to note that the simulations in this catalogue do not incorporate active galactic nucleus (AGN) feedback. As a result, the galaxies formed in these simulations exhibit an excessive mass and an ultradense stellar distribution at low redshifts (Çatmabacak et al. 2022).



**Figure 2.1** – Density profile of the A1 simulated halo from the MassiveFIRE suite of simulations. The SOI of the SMBH at the center of this halo is shown by the red circle. Here, the SMBH is around  $10^9 M_{\odot}$  which leads to a radius of  $\sim 150pc$ . We see the relatively small size of the SOI compared to the halo.

The MassiveFIRE suite consists of four halos (A1, A2, A4, A8), initially selected from the A-series of the FIRE-1 MassiveFIRE suite (Feldmann et al. 2016), which enables us to explore a wide range of formation histories for halos with a typical virial mass of  $M_{\text{vir}} \approx 10^{12.5} M_{\odot}$  at  $z = 2$ . For details regarding the selection strategy and halo growth in the MassiveFIRE catalogue, we refer to Feldmann et al. 2017.

The MassiveFIRE catalogue provides 19 full snapshots for each halo, spanning from redshift 1 to 10 with a redshift interval of  $\delta z = 0.5$  (Anglés-Alcázar et al. 2017b). In our analysis, we utilize all available halos, amounting to over 400 snapshots within our region of interest, specifically the redshift range from 1 to 2. This redshift range represents the region where most lensed galaxies are located. These snapshots offer valuable insights into the evolution and properties of the galaxies and their associated SMBH during this critical period.

## 2.3. Simulations

### 2.3.1. Black Hole injection

In our study, we work with four independent galaxies observed at multiple snapshots spanning from redshift 1 to 2. Each snapshot represents a distinct data point since the mass of the SMBH varies over time. To further enhance the quantity and diversity of mock ALMA cubes generated (detailed in Section 2.3.4), we introduce manual variations in the SMBH

mass, ranging uniformly from  $10^6$  to  $10^{10}M_{\odot}$ . By comparing the original SMBH mass to the new values, we can calculate the gain or loss of potential energy at various distances from the SMBH. Assuming a virialized system, we then establish a correlation between potential energy and velocity. Finally, we scale these values based on the initial potential energy (prior to the variation of SMBH mass), enabling a direct calculation of the final velocity,  $v_f$ , based on the initial velocity,  $v_i$ , initial potential,  $U_i$ , and final,  $U_f$ . More precisely :

$$U = -2T, \tag{2.3.1}$$

$$U = -2\left(\frac{1}{2}mv^2\right), \tag{2.3.2}$$

$$U = -mv^2, \tag{2.3.3}$$

$$\frac{U_i}{U_f} = \left(\frac{v_i}{v_f}\right)^2, \tag{2.3.4}$$

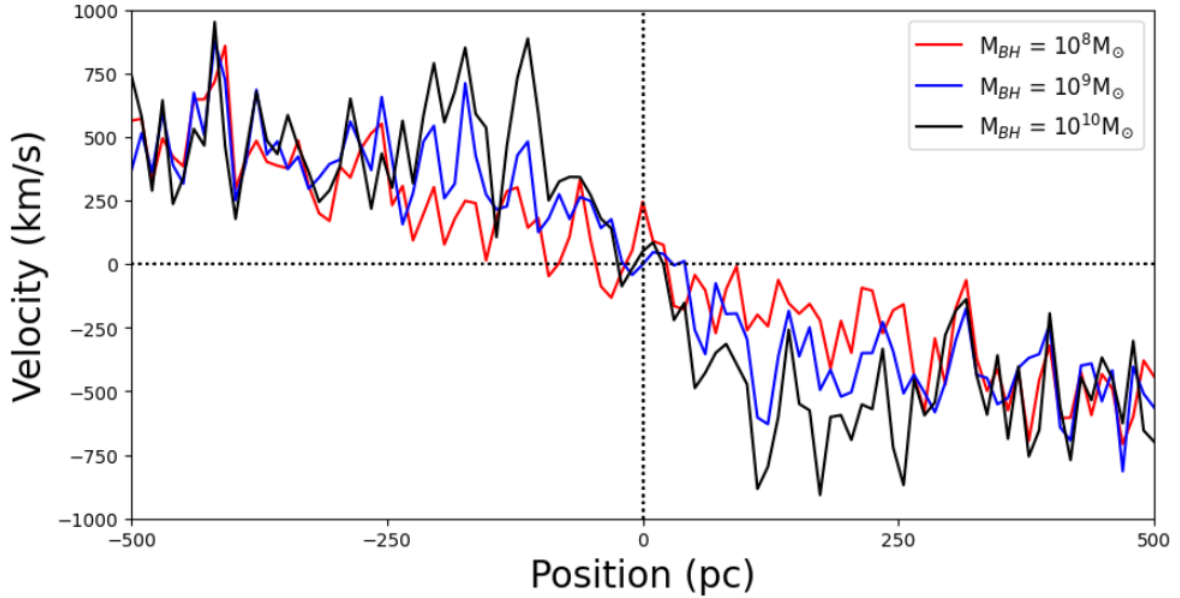
$$v_f = v_i\sqrt{\left(\frac{U_f}{U_i}\right)}, \tag{2.3.5}$$

hence, the velocity of the molecular gas in each snapshots is rescaled based on the newly injected SMBH mass.

It is important to note that the sudden change in SMBH mass would have various effects on the immediate environment, extending beyond purely kinematic changes. Specifically, we would expect the temperature and density of the molecular gas to increase. The higher dynamics resulting from the SMBH mass change would lead to more energy in collisions, resulting in increased heat emission. Additionally, a more massive SMBH would create a deeper potential well, leading to a more condensed molecular gas distribution. Over time, this change in SMBH mass would also impact the overall morphology and evolutionary history of the host galaxy (Vogelsberger et al. 2014). However, for the purpose of this study, where we aim to train a deep neural network to detect the kinematic effects of the SMBH, altering the dynamics of the molecular gas alone should be sufficient. In fact, this approach is beneficial for the interpretability of the network. By focusing solely on the gas kinematics, which provide the most accurate and direct information for estimating the SMBH mass, we can

have greater confidence in the network’s ability to discern the relevant kinematic signatures. This allows us to avoid potential confounding factors such as the flux (which depends on other physical parameters like temperature and density) and the galaxy morphology, which are also known to correlate with the mass of the SMBH (Dunlop et al. 2003).

For each injected SMBH, we can compute the velocity in the line of sight as a function of distance from the SMBH. Figure 2.2 highlights this for the same galaxy injected with three different SMBH masses. As the mass of the SMBH increases, the slope of the initial Keplerian velocity rise becomes increasingly steep. However, it is crucial to note that at significant distances from the SMBH, the velocities approach degeneracy. As anticipated, this occurs because the gravitational effect of the SMBH is no longer dominating over the effect of the galaxy.



**Figure 2.2** – Line-of-sight velocity (km/s) as a function of distance from the SMBH. The three distinct curves correspond to three different SMBH masses implemented in the same galaxy. The SOIs for the SMBHs of masses  $10^8 M_{\odot}$ ,  $10^9 M_{\odot}$ ,  $10^{10} M_{\odot}$  are respectively roughly 50, 75, 150 parsecs. We see that at shorter distances, the greater the SMBH mass, the higher the absolute velocity. However, as we start to approach distances beyond the SOI, the differences in velocities progressively disappear.

### 2.3.2. CO emission line

Our goal is to generate a simulated ALMA data cube, specifically observing the CO(2-1) emission line because it has been shown to be the best molecular line tracer for SMBH mass measurements (Davis et al. 2017). This process begins with the identification of molecular gas within the FIRE-2 simulation. These snapshots encompass four distinct types of matter : dark matter, stars, gas, and black holes. The gaseous component within the simulation

displays a broad range of properties, with temperatures ranging from 10 K to  $10^{10}$  K and densities ranging from  $10 \text{ cm}^{-3}$  to  $10^5 \text{ cm}^{-3}$ . To define and isolate the molecular gas within this context, specific criteria are applied. Here, we consider gas with a temperature lower than  $10^3$  K to be classified as molecular gas (Hopkins et al. 2014a).

To generate mock ALMA data, we establish a consistent and accurate methodology for converting the physical parameters of the simulated CO molecular gas into emitted flux values. In this regard, we provide an overview of the flux calculator methodology, while directing readers to Weiss et al. *in prep.* for a detailed, step-by-step description of the method. For a chosen emission line with frequency,  $\nu$ , the flux density of a cloud of gas,  $S_\nu dv$ , is obtained using

$$S_\nu dv = \frac{I_{\text{line}} \cdot dv}{1+z} \cdot \frac{A_{\text{size}}}{D_A^2} \cdot \frac{2k \left(\frac{\nu}{1+z}\right)^2}{10^{-26} c^2}. \quad (2.3.6)$$

Here,  $I_{\text{line}}$  is the line intensity,  $dv$  is the velocity dispersion,  $z$  is the redshift,  $A_{\text{size}}$  is the size of the area,  $D_A$  is the angular size distance,  $k$  is the Boltzmann constant,  $\nu$  is the frequency of the emission line, and  $c$  is the speed of light. Except for  $I_{\text{line}}$ , every other parameter in Equation 2.3.6 can directly be obtained from the simulation.

To calculate  $I_{\text{line}}$ , the radiative transfer equation has to be solved. This equation describes how the intensity of radiation changes as it propagates through the gas. For a beam of radiation with specific intensity  $I_\nu$  traveling through a medium with absorption coefficient  $\alpha_\nu$  and emission coefficient  $\epsilon_\nu$ , the radiative transfer equation is :

$$\frac{dI_\nu}{ds} = -\alpha_\nu I_\nu + \epsilon_\nu. \quad (2.3.7)$$

Equation 2.3.7 states that the change in the specific intensity along the path,  $ds$ , is due to absorption (which decreases the intensity) and emission (which increases the intensity). In the context of line emission from molecular gas, the absorption and emission coefficients are determined by the level populations and the Einstein coefficients for the relevant transitions. By assuming statistical equilibrium, for a given energy level,  $i$ , the rate of change of the population,  $n_i$ , is given by :



$$\frac{dn_i}{dt} = \sum_j (n_j A_{ji} + n_j B_{ji} J_{\nu_{ji}} + n_j C_{ji} - n_i A_{ij} - n_i B_{ij} J_{\nu_{ij}} - n_i C_{ij}) = 0, \quad (2.3.8)$$

where  $n_j$  are the populations of the other levels,  $A_{ji}$ ,  $B_{ji}$ , and  $C_{ji}$  are the Einstein coefficients for spontaneous emission, stimulated emission, and absorption, respectively, between levels  $j$  and  $i$ , and  $J_{\nu_{ji}}$  is the mean intensity of the radiation field at the frequency of the transition between levels  $j$  and  $i$ . The statistical equilibrium assumption is that the populations do not change with time, so the rate of change is set equal to zero.

To solve for the line intensities, Equation 2.3.7 and 2.3.8 need to be solved together, because the level populations determine the absorption and emission coefficients, which affect the intensity of the radiation, which in turn affects the level populations through  $J_{\nu_{ji}}$ . The initial values for  $I_\nu$  and  $n_i$  are set by the temperature, density and  $H_2$  fraction of the gas cloud, all of which can be directly obtained from the simulation.

The simultaneous solution of this system of equations is facilitated by the assumption of a turbulent large-velocity-gradient (LVG) model. This model posits that the velocity of the gas changes significantly over the size of each emitting region, and each region is only in radiative contact with itself. This simplifies the radiative transfer equation by rendering it local in nature. As a consequence, the collision rates  $C_{ji}$  are assumed to depend solely on the local properties of the gas, and the radiation field  $J_{\nu_{ji}}$  is assumed to be constant throughout the gas.

Beginning with the initial values of  $I_\nu$  and  $n_i$ , the simultaneous solution of the radiative transfer and statistical equilibrium equations (Equations 2.3.7 and 2.3.8 respectively), until convergence, provides the evolution of the intensity as it propagates through the molecular gas for a chosen frequency. This enables the calculation of  $I_{line}$ , the last missing variable needed to solve Equation 2.3.6. As a result, we have an automated method to calculate the emitted flux of a molecular gas region. This flux corresponds to a CO transition line ranging from CO(1-0) to CO(15-14). For this study, we will focus on the CO(2-1), which has been demonstrated to be one of the most suitable molecular lines for measuring the mass of SMBHs using molecular gas dynamics (Davis et al. 2011). Given that each region of the molecular gas is treated as independent due to the LVG assumption, this task can be efficiently parallelized, leading to a significantly faster computation.

### 2.3.3. Gravitational Lensing

Our forward model accepts a mock ALMA observation of a galaxy as input and generates a distorted representation of the observed galaxy, employing a strong gravitational lens. Strong gravitational lensing transpires when an observed light ray, emitted from a background

source, is bent by a massive foreground object. The lens equation relates the angular position of the source in the source plane,  $\vec{\beta}$ , to the lensed angular position in the lens plane  $\vec{\theta}$  :

$$\vec{\beta} = \vec{\theta} - \vec{\alpha}(\vec{\theta}), \quad (2.3.9)$$

where  $\vec{\alpha}$  signifies the deflection angles induced by the deflector. The potential of the massive object causing the deflection, denoted by  $\phi$ , dictates  $\vec{\alpha}$  as follows :

$$\vec{\alpha}(\vec{\theta}) = \vec{\nabla}\phi(\vec{\theta}). \quad (2.3.10)$$

This potential depends on the dimensionless surface mass density,  $\kappa$ , also known as the convergence map :

$$\nabla^2\phi(\vec{\theta}) = 2\kappa(\vec{\theta}). \quad (2.3.11)$$

From Equations 2.3.10 and 2.3.11, the following relationship between  $\vec{\alpha}$  and  $\kappa$  is derived :

$$\vec{\alpha}(\vec{\theta}) = \frac{1}{\pi} \int_{R^2} \frac{\vec{\theta} - \vec{\theta}'}{|\vec{\theta} + \vec{\theta}'|^2} \kappa(\vec{\theta}') d^2\theta'. \quad (2.3.12)$$

The convergence map is then determined by the entire lensing system :

$$\kappa(\vec{\theta}) \equiv \frac{\Sigma(\vec{\theta})}{\Sigma_c(\vec{\theta})} = \frac{4\pi G}{c^2} \frac{D_L D_{LS}}{D_S} \Sigma(\vec{\theta}). \quad (2.3.13)$$

Here,  $\Sigma_c(\vec{\theta})$  represents the critical density, a characteristic feature of the lensing system that depends on the angular distance between observer and lens plane ( $D_L$ ), observer and

source plane ( $D_S$ ), and lens and source plane ( $D_{LS}$ ). The surface density,  $\Sigma(\vec{\theta})$ , is a function of the foreground's mass profile.

In this study, we adopt the singular isothermal ellipsoid (SIE) mass profile, which depends on five parameters. Firstly, the central position of the deflector in both x and y directions must be established. Secondly, the eccentricity of the matter distribution of the deflector in x and y needs to be determined. The last parameter is the Einstein radius defined as :

$$\theta_E = \frac{2}{c} \sqrt{GM(\theta_E) \frac{D_{LS}}{D_L D_S}}, \quad (2.3.14)$$

where G is the gravitational constant and  $M(\theta_E)$  is the total mass of the deflector enclosed within the Einstein radius. For practical purposes, due to the complete degeneracy between mass and distances, we fix the redshifts of the main deflector at  $z = 0.5$  and vary only the total mass of the lensing galaxy to adjust the Einstein radius. The source's redshift is already embedded when generating the mock ALMA data. Utilizing the SIE profile enables us to analytically calculate the magnification index at each pixel based on these five parameters, affording us the flexibility to select specific parameter combinations to attain desired levels of magnification for accurately measuring the mass of a SMBH.

The lens equation (Equation 2.3.9) can be employed to obtain the position in the source plane given the position in the lens plane, but not vice versa. Therefore, to simulate the distortion of a source, we first need to pixelate the lens plane and, using the lens equation, calculate the originating position of that pixel in the source plane. From this pixel-to-pixel correlation of the lens and source planes, we can map the source to the lens plane using a bilinear interpolation. Importantly, gravitational lensing is an achromatic phenomenon, implying it is independent of the color of light. As a result, we can efficiently parallelize the lensing process across the different generated ALMA cubes and their channels.

### 2.3.4. Mock ALMA data

The first step in simulating ALMA observations involves generating an empty data cube by pixelating the simulations both spatially and spectrally. To achieve the highest spatial resolution, we consider the most extended configuration for ALMA with baselines up to 16 km. For the CO(2-1) emission line at a frequency of  $f_{\text{CO}(2-1)} = 230.5$  GHz, the spatial resolution is 20 milliarcseconds (mas). The field of view is determined by the frequency and configuration, typically defined by the full-width-at-half-maximum (FWHM) of the primary beam. In the chosen configuration of ALMA, the FWHM for CO(2-1) is 15 arcseconds. With this information, we construct a 2D-array representing the X and Y positions in the sky.

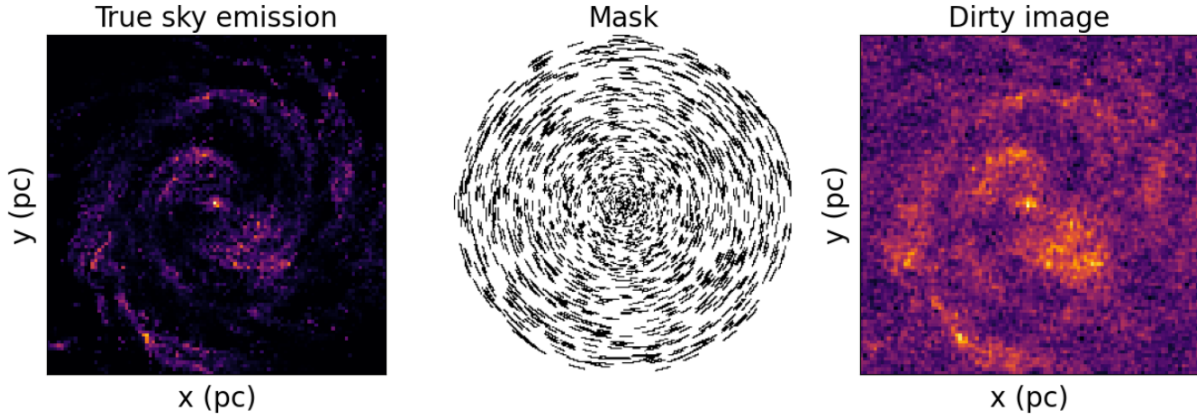
For the spectral resolution, ALMA can deliver data cubes with up to 7680 frequency channels. The width of these channels can range from 3.8 kHz to 15.6 MHz, but the total bandwidth cannot exceed 8 GHz. To ensure that the spectral dimension does not limit our analysis, we use the highest possible spectral resolution for the CO(2-1) line at  $f_{\text{CO}(2-1)}$ . This yields a velocity resolution of approximately 10 km/s, which is adequate to discern the influence of a SMBH provided that the SOI is resolved. We utilize 100 channels, culminating in a bandwidth of 1 GHz. This bandwidth permits us to span velocities from -500 km/s to 500 km/s. Such a range is sufficient to thoroughly characterize the kinematic effects of a SMBH on CO molecular gas, given that even molecular gas surrounding the most massive SMBHs barely surpasses this velocity span (Ruffa et al. 2023). With these specifications, we create an empty cube for our mock ALMA data and then populate it with relevant information.

For every simulated observation, we select a specific 2D projection. However, to create a comprehensive data cube, we also establish a spectral distribution for each pixel along the line of sight. This involves considering the velocity distribution of the molecular gas within each pixel. Starting with the velocity distribution, we discretize it based on the predefined spectral resolution. It is important to note that velocities and frequencies are interchangeable using the Doppler shift given by  $f = \left(1 + \frac{v}{c}\right) f_{\text{CO}(2-1)}$ . This allows us to pixelate the distribution and assign a velocity value to each spectral channel. By repeating this process for every velocity channel, we obtain a velocity cube that accurately represents the distribution of velocities in the line of sight at every position.

Next, we utilize the same Doppler shift equation to convert the velocity channels ( $v$ ) into corresponding frequencies ( $f$ ). This transformation takes into account the frequency shift caused by the relative motion of the gas in relation to the observer. By applying this equation to each velocity channel, we effectively translate the velocity information into the frequency domain. Additionally, we consider the frequency shift resulting from the expansion of the Universe. In this case, since the spatial distances of the molecular gas along the line of sight are negligible on cosmological scales, this shift corresponds to a simple translation in the frequency domain, dependent only on the redshift of the galaxy.

The frequency space of the data now has to be renormalized based on the modeled observed CO(2-1) integrated flux. To accomplish this, we calculate the integrated flux for each pixel in the 2D projection using the previously discussed method. This integrated flux map provides the total flux emitted by each pixel across the spectral dimension. We use this map to renormalize every pixel in the spectral dimension so that the values in every pixel no longer represent a count but a flux. This adjustment entails modifying the values in the spectral dimension so that the sum of flux within each pixel corresponds to the integrated flux emitted from the specific region in the 2D projection. Through this renormalization process, we ensure that the flux distribution across the spectral dimension accurately represents the emission characteristics of the simulated patch in the 2D projection. This step guarantees

a properly calibrated and representative data cube that aligns with the observed physical properties of the molecular gas in the simulation.

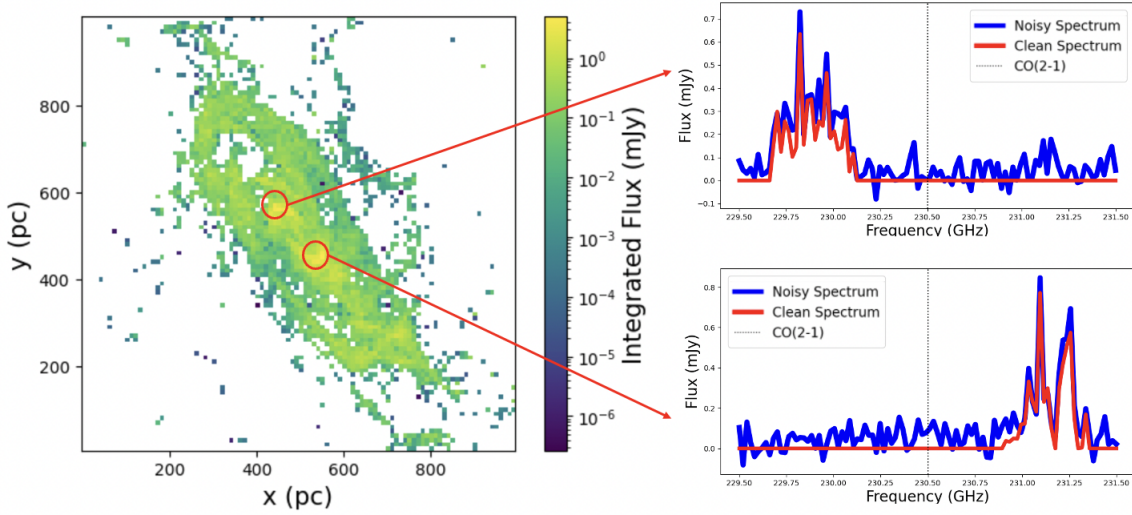


**Figure 2.3** – The impact of the masked Fourier transform on a clean image is depicted in the following sequence. On the left, we present a flattened representation of a simulated ALMA cube. This image is the true sky emission, devoid of noise or artifacts. In the middle, we display a typical mask employed for sampling from the U-V plane; black represents the measured fourier nodes, while the white are the unobserved. The dirty image is obtained by conducting an inverse Fourier transform on the masked frequency domain, to which Gaussian noise has been added.

Given that ALMA is an interferometer, the generated cubes have to undergo a Fourier transformation into the frequency domain. We then add Gaussian noise to the cubes to reflect the inherent nuisance in observational data. To transition back to the spatial domain, a mask corresponding to the measured points in the U-V plane must be generated. This mask is determined by the number of antennas, their configuration and the exposure time. The number of antennas directly influences the quantity of baselines which, in essence, probe the U-V plane. The configuration determines the extent of the U-V plane being explored. The exposure time, in conjunction with Earth’s rotation, probes additional areas within the U-V plane. This is why the generated mask exhibits a circular pattern (in Figure 2.3). Here, our aim is to examine a maximal area in the U-V plane, while obtaining the highest possible spatial resolution. Consequently, we deploy all 66 of ALMA’s antennas, yielding 4290 baselines, in the D configuration for an exposure time of 4 hours. It’s important to note that a 4-hour duration is optimal for comprehensive exploration of the U-V plane, facilitated by Earth’s rotation. Once the mask is chosen, we can apply it to the full U-V plane and revert to the spatial domain via an inverse Fourier transform. This final output is known as the ‘dirty image’, a process which is replicated across all channels.

With our completed pipeline, we can rapidly generate a simulated ALMA cube, incorporating the kinematical effect of a SMBH of an arbitrary mass. Figure 2.4 provides a representative example of such a cube. On the left, we display a projection of the cube, derived by

summing all channel information. The right side presents spectra at two different positions within the cube, adjusted for redshift. This figure shows the displacement of the CO(2-1) emission line. From this shift, one can not only discern the overall rotational direction of the galaxy but also estimate the mass of the supermassive black hole located at its center.



**Figure 2.4** – Representation of a generated mock ALMA cube observing the CO(2-1) emission from molecular gas. On the left, we see the integrated spectra of the cube in units of mJy. On the right, we have the spectrum at 2 different pixels located at the center of the red circles. Considering the spectral shift of the emission from the CO(2-1) rest frequency at both locations, we can infer the axis of rotation of the entire galaxy. The methodology for kinematically measuring the mass of SMBHs follows a similar procedure, but focused within the SOI.

### 2.3.5. Data augmentation

Data augmentation plays a crucial role in machine learning, especially considering the data-intensive nature of neural network training. All the simulations are conducted in three-dimensional space. However, in practice we only observe a two-dimensional projection of this space along a specific line of sight. By uniformly sampling from a range of projection angles ( $\theta$ ), we can further augment the dataset and introduce variations in the observed orientations of the galaxies. It is important to note that all these projections maintain the same SMBH mass, allowing the network to learn the ability to generalize across different viewing angles. In addition to the projection angle, we also consider the inclination angle ( $\phi$ ) of the galaxy, which has a significant impact on the network’s capacity to accurately estimate the SMBH mass. When  $\theta = 0$ , signifying a face-on perspective, there is an absence of kinematic information along the line of sight. This lack of information makes it exceptionally challenging for the network to accurately predict the mass of the SMBH. To address this, our dataset includes a uniform distribution of inclinations ranging from  $\pi/6$  to  $\pi/2$  (edge-on).

Lastly, we further augment the dataset by applying basic image transformations, aiming to train a network that is invariant to non-relevant changes in the image. This generative pipeline is fully automated. We generate a mock ALMA data cube as a function of redshift, SMBH mass and inclination ( $\phi$ ).

## 2.4. Neural Network

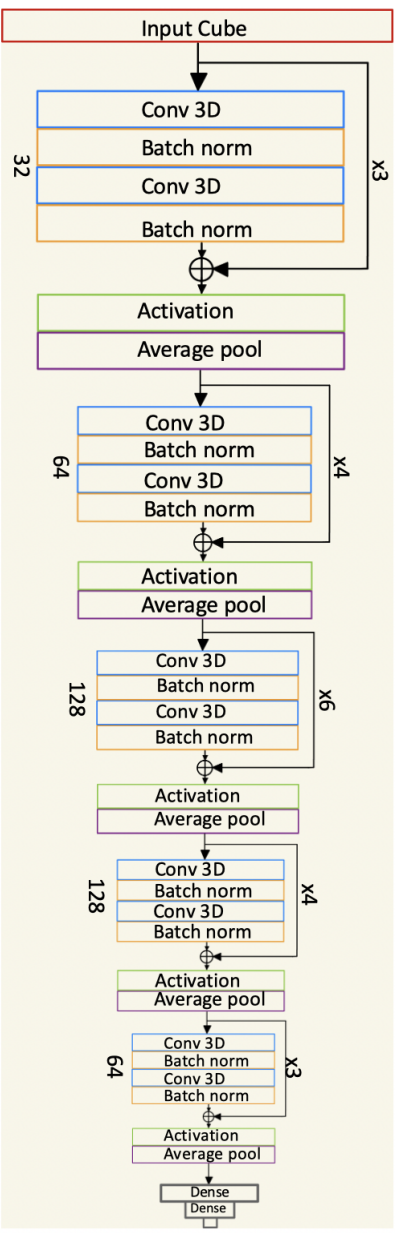
Traditional methods for SMBH mass measurements, such as Markov Chain Monte Carlo (MCMC) sampling, rely on approximations based on assumptions of symmetry and circular orbits (Boizelle et al. 2019). However, these methods can be time-consuming and may overlook the intricate spatial and spectral correlations inherent to black hole phenomena. In this study, we propose an alternative approach that harnesses the power of Convolutional Neural Networks (CNNs) to rapidly and accurately estimate the mass of SMBHs without the need for such approximations.

### 2.4.1. Architecture

The architecture that we use is a deep neural network model that belongs to the family of Residual Networks (ResNets). ResNets were introduced to address the vanishing gradient problem that often occurs in very deep neural networks (Szegedy et al. 2016). The vanishing gradient problem refers to the issue of gradients diminishing as they propagate through multiple layers, making it difficult for the network to learn effectively. When adding multiple hidden layers to a network, we augment its expressivity, making it possible to learn more complex non-linear correlation dependent on multiple parameters. That being said, when backward propagating a gradient through a lot of layers, we drastically augment our odds of vanishing gradient. The key innovation of ResNets is the use of residual connections (Shrivastava et al. 2016). These connections allow the network to learn residual mappings, which are the differences between the desired output and the current output of a layer. By propagating the residuals directly through the network, the gradients have a shorter path to flow, making it easier for the network to learn and allowing efficient training of deeper networks.

Our network architecture consists of 40 convolutional layers, plus normalization, pooling and fully connected layers (see Figure 2.5). It follows a *building block* structure, where each building block consists of two convolutional layers after which comes a batch normalization (Ioffe & Szegedy 2015). Once the block is exited, the information is down sampled via average pooling and goes through an *elu* activation function. The residual connections bypass these convolutional layers and add the original input to the output of the block. This allows the network to retain important information from earlier layers and facilitate gradient flow. The residual connections aid in feature extraction. They enable the network to capture and learn

both low-level and high-level features by preserving information from earlier layers (LeCun et al. 2015). This capability is particularly beneficial for complex tasks such as regression. Also, we note that since the gradient flow is improved, the network can converge faster, leading to shorter training times and reduced computational resources.



**Figure 2.5** – Architecture of the CNN trained to measure the mass of SMBHs. The network is formed of various 3D convolution building blocks with a different channel size (numbers at the bottom) that are repeated multiple times (numbers at the top). The input cube is size  $96 \times 96 \times 96$  and is downsampled by average pooling until no more spatial correlation is left. At that time, the cube is flatten and is passed through a couple a dense layers to finally output two values : the SMBH mass and the standard deviation.

CNNs are powerful tools for analyzing spatially correlated data, making them well-suited for accurately measuring the mass of a SMBH through molecular gas kinematics. To achieve this, we must consider both spatial and spectral correlations. Consequently, our model utilizes three-dimensional convolutional layers, enabling an efficient capture of the kinematic effects of the SMBH.

When estimating the mass of a SMBH based on kinematics, it is necessary to fit the velocity as a function of the radial distance from the SMBH. As we move farther away from the SMBH, the amount of matter within the orbit increases, resulting in a corresponding increase in the absolute value of the velocity. The mass of the SMBH is thus an inherent parameter of this rise. To enable the network to perform this fitting process internally, it requires the ability to access spatial correlations between position and spectrum. This is precisely why we introduce 3D convolutional layers into our model, as they provide the network with the necessary tools to effectively correlate position and velocity information.

The pooling layers are used to down sample the data, in our case, until we have only 1 value, the SMBH mass. Average pooling and max pooling are two common techniques used in CNNs for downsampling feature maps. While both pooling methods serve the purpose of reducing the spatial dimensionality of feature maps, they differ in how they aggregate information within pooling regions.



To ensure the optimal performance of our model, we carefully tune the hyperparameters. Two critical hyperparameters we consider are the batch size and learning rate. The batch size is limited by GPU memory. Larger batch sizes can lead to faster convergence during training (You et al. 2017). By processing more examples per iteration, the model can update its weights more accurately and make more significant progress in each update. This can result in faster training and require fewer iterations to achieve a certain level of performance. Using larger batch sizes can have a regularization effect on the model. In practice, larger batch sizes tend to generalize better and can help reduce overfitting. This regularization effect is attributed to the fact that the model is exposed to more diverse examples within each batch, forcing it to learn more robust and generalized representations (Lei Ba et al. 2016). In our case, because we are dealing with data cube that take a lot of memory space, we go for a batch size for 32, the maximum amount of cube that fit in our GPU’s memory. As for the learning rate, through trial and error optimization we found that for our problem, an exponentially decaying learning rate starting at  $10^{-4}$  leads to a fastest convergence and more accurate predictions. Finally, we work with the stochastic optimization *Adam* to optimizer our weights during training (Kingma & Ba 2014).

## 2.4.2. Estimating Gaussian Uncertainties

Accurate estimation of uncertainties is a crucial component of regression tasks. Alongside obtaining point predictions, it is essential to assess the associated uncertainty for each prediction. To address this, we introduce a Gaussian loss function that directly models uncertainties by assuming a Gaussian posterior distribution (Zou et al. 2018).

The Gaussian loss function is defined as follows :

$$L = -\log \left[ \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left( -\frac{(y_{\text{pred}} - y_{\text{true}})^2}{2\sigma^2} \right) \right], \quad (2.4.1)$$

$$= \frac{1}{2} \left[ \log(\pi\sigma^2) + \frac{(y_{\text{pred}} - y_{\text{true}})^2}{\sigma^2} \right]. \quad (2.4.2)$$

Here,  $y_{\text{pred}}$  represents the network’s predicted value,  $y_{\text{true}}$  denotes the true value, and  $\sigma$  represents the standard deviation of the predicted value, assuming a Gaussian distribution.

This Gaussian loss function can be interpreted as the negative log-likelihood of the true SMBH mass given the predicted mean and standard deviation. By employing this loss function, we capture the discrepancy between the predicted distribution and the true target value,

encouraging the model to learn both the point estimates and the uncertainty associated with each prediction. The parameter  $\sigma$  is unsupervised and self-regulating, as it is directly and inversely proportional to the loss. Minimizing the Gaussian loss allows the model to adapt both the mean and standard deviation of the predicted Gaussian distribution. Consequently, the model becomes capable of capturing the inherent uncertainty present in the data and providing more accurate estimates of the uncertainties.

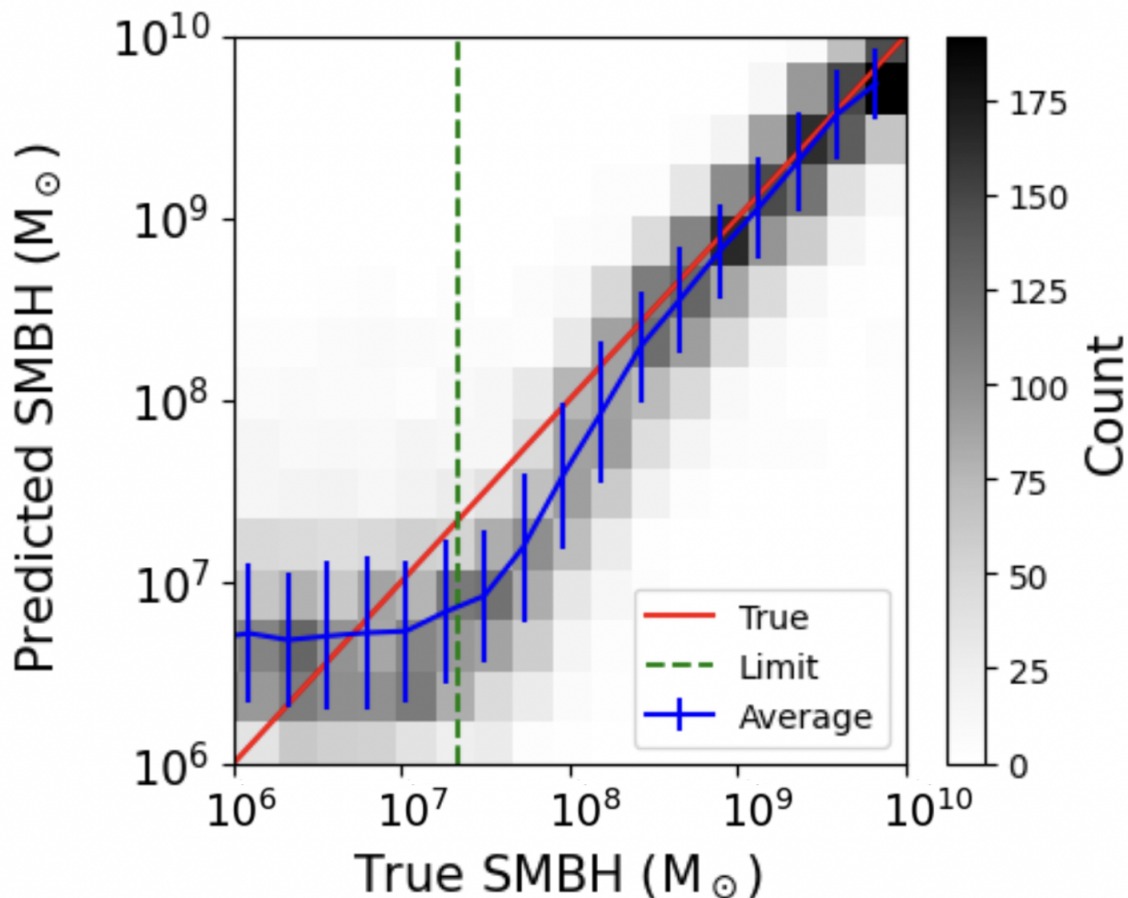
Using a Gaussian loss for uncertainty estimation offers several advantages, including simplicity and ease of implementation. The loss function can be readily incorporated into existing machine learning frameworks, and the model can be trained using standard optimization algorithms such as stochastic gradient descent. However, it is important to acknowledge that the effectiveness of the Gaussian loss approach relies on the assumption that the true uncertainties follow a Gaussian distribution. While this assumption may not always hold in practice, it provides a reasonable approximation for estimating the uncertainties associated with SMBH mass measurements.

## 2.5. Discussion

The fully developed generative pipeline is employed to assess the feasibility of predicting the mass of SMBHs from realistic, simulated ALMA data. Figure 2.6 illustrates the predictive performance of the network trained on a dataset of over 100,000 cubes containing uniformly distributed galaxies from the MassiveFIRE suite of massive halos. These cubes cover a redshift range from  $z = 1$  to  $z = 2$ , with inclinations varying from 0 (face-on) to  $\frac{\pi}{2}$  (edge-on), and SMBH masses ranging from  $10^6 M_\odot$  to  $10^{10} M_\odot$ . The spatial resolution is set at 10 pc per pixel, and the spectra resolution is 10 MHz per channel. The results indicate a high level of accuracy in predicting SMBH masses for masses above  $10^8 M_\odot$ . This outcome aligns with the expectation that heavier SMBHs significantly impact the kinematics of molecular gas.

For lower masses, the network tends to regress to the average value within the remaining SMBH mass range. This behavior is common in machine learning, where, when unable to provide an accurate prediction, the network minimizes loss by outputting the average value. However, in our case, this behavior further validates that the network relies on kinematic information rather than other hidden features in the data to estimate SMBH masses.

Indeed, this method of SMBH mass measurements is effective only when the SOI of a SMBH is resolved. Since a spatial resolution of 10 pc per pixel is not enough to resolve SOI of a SMBH with a mass smaller than  $10^8 M_\odot$ , we anticipate the network to fail in such cases. We conducted multiple training iterations at various spatial resolutions, and the observed plateau in performance corresponds to the aforementioned correlation. Specifically, higher

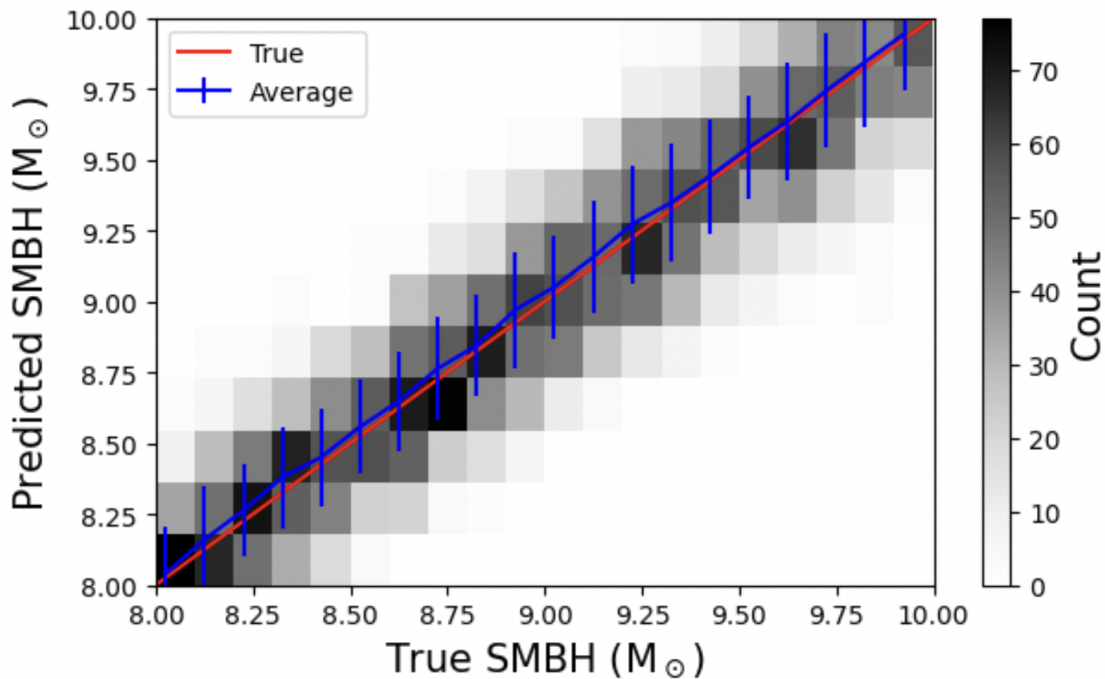


**Figure 2.6** – Predicted SMBH mass ( $M_{\odot}$ ) of the network as a function of the true SMBH mass ( $M_{\odot}$ ). The red line represents the ideal prediction. In black are the binned predictions with the mean and standard deviation in blue. The green vertical line represents the critical SMBH mass where the physical scale of a pixel is the same size as the SOI. We see that for masses higher than that critical line, the predictions are accurate. For smaller masses, the spatial resolution is no longer good enough to resolve the SOI and the network loses the ability to accurately predict the mass of those SMBHs.

spatial resolutions enable the network to accurately predict the masses of SMBHs at lower mass ranges, coinciding with the resolvability of the SOI.

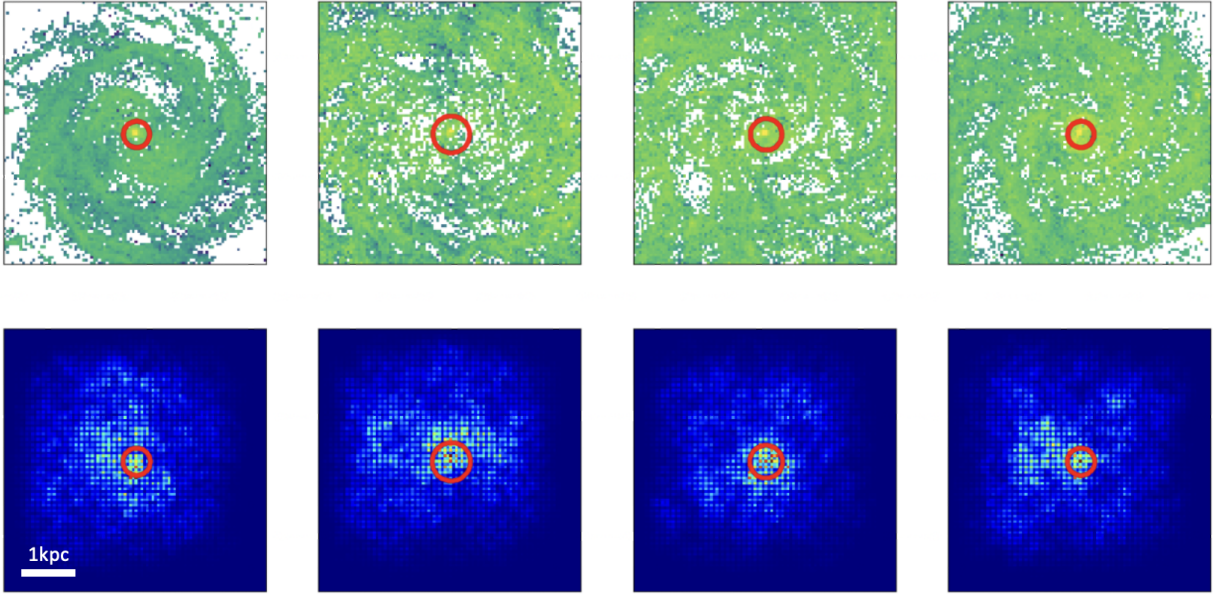
A noteworthy source of bias arises when the network is tasked with predicting the mass of a SMBH within a range of masses transitioning from unresolved to resolved SOI. This leads to a significant underestimation of mid-range SMBH masses, despite the presence of information on the kinematical effect of the SMBH. We mitigated this problem by training on data known to have sufficient spatial resolution to resolve the SOI, thereby avoiding bias in our predictions (see Figure 2.7). This effect is also observable in the inferred uncertainties : predictions are very confident at higher masses and progressively become more uncertain with lower SMBH masses. This aligns with the notion that heavier SMBHs possess larger

SOI, making their mass inference easier. However, when the network is trained solely on a range of high SMBH masses, all predictions reach the same level of high confidence.



**Figure 2.7** – Predicted SMBH mass ( $M_{\odot}$ ) of the network as a function of the true SMBH mass ( $M_{\odot}$ ). Here, the network is only trained and tested on SMBH with high enough masses for the SOI to be resolved considering the spatial resolution. Contrary to Figure 2.6, there is no bias in any of the predictions.

In an effort to confirm that the network is directing its primary focus towards the SOI when predicting SMBH masses - as opposed to exploiting other features in the data such as the size of the galaxy or its total flux - we conducted an analysis of saliency maps derived from several examples (refer to Figure 2.8). Saliency maps, generated by backpropagating the gradient from the network’s prediction to the original input data, help pinpoint the pixels that carry the greatest influence over the prediction. They provide valuable insights into the decision-making mechanisms of our neural network and serve as a representation of the network’s focal points, accentuating the critical regions in the data that contribute most significantly to the final prediction. Figure 2.8 shows the network’s ability to concentrate its attention on the SOI during the SMBH mass prediction. This is a crucial finding, given the SOI’s role as the dominant region for the gravitational influence of the SMBH. However, it’s important to consider that while these saliency maps validate the network’s focus on the anticipated areas, they do not offer comprehensive insights into the internal reasoning of the network, nor do they expose any potentially complex relationships that the network may have discerned from the data.



**Figure 2.8** – Saliency maps corresponding to accurately predicted SMBH masses from a variety of galaxies. The red circle represents the SOI. The top row displays the projected mock ALMA cubes from four distinct galaxies, while the bottom row presents their respective saliency maps. These maps are obtained by backpropagation through the network, tracing from the prediction to the input data. The prominently highlighted pixels in the saliency maps correspond to the SOI, confirming that this region, which contains the most crucial information for accurately measuring the SMBH mass, is indeed driving the network’s predictions.

## 2.6. Conclusion

Prior research has explored the potential of leveraging gravitational lensing to resolve the sphere of influence (SOI) of supermassive black holes (SMBHs) at higher redshifts. For instance, Hezaveh 2014 demonstrated the feasibility of this approach with a basic model. In this study, we put forth a more sophisticated method to simulate the impact of a SMBH on its surrounding kinematics. Our pipeline is designed to rapidly generate Atacama Large Millimeter/submillimeter Array (ALMA) data cubes, thereby enabling the training of data-intensive methodologies for precise SMBH mass measurements.

We developed a rapid, automated pipeline to generate mock ALMA data cubes for inferring SMBH masses. Beginning with the MassiveFIRE hydrodynamical simulation catalogue, we initially integrated the kinematic effects of a SMBH based on its mass, presuming a virialized system. From these simulations, we calculate the integrated CO(2-1) emission line using parameters such as velocity,  $H_2$  fraction, density, temperature, distance, and the size of the observed sky region. Subsequently, the mock ALMA cube is translated into the frequency domain, where Gaussian noise is introduced and a mask is applied in accordance with

ALMA’s configuration, bandlines, and exposure time. We test the performance of our generative pipeline using a regression task to estimate SMBH masses and their uncertainties via a three-dimensional convolutional neural network (CNN), assuming a Gaussian posterior. Our findings align with established methods, affirming that accurate SMBH mass measurements can be obtained when the SOI is adequately resolved.

With the forthcoming operation of Euclid and LSST, over 200,000 gravitational lensing systems are projected to be uncovered. Consequently, our future research will focus on the exploration of lensed mock ALMA data to investigate the possibility of measuring SMBH masses at elevated redshifts. Our approach will entail reconstructing the background galaxy along with all its kinematic effects using a score-based model approach.

## 2.7. Acknowledgements

D.C. acknowledges financial support from NSERC, FRQNT and IVADO via excellence scholarships, as well as from the physics department of the Université de Montréal and the Center of Research in Astronomy of Quebec. The work is in part supported by computational resources provided by Calcul Quebec, Compute Canada and the Digital Research Alliance of Canada. Y.H. and L.P. acknowledge support from the National Sciences and Engineering Council of Canada grant RGPIN2020-05102, the Fonds de recherche du Québec grant 2022-NC301305 and 300397, and the Canada Research Chairs Program. J. H.-L. acknowledges support from NSERC via the Discovery grant program, as well as the Canada Research Chair program.

# Chapitre 3

---

## Conclusions

Tel qu'introduit dans ce mémoire, les trous noirs supermassifs (SMBH) présents au sein des galaxies sont des objets caractérisés par une gravité si intense que même la lumière ne peut échapper à leur puits gravitationnel. De manière surprenante, la masse de ces SMBHs présente une corrélation avec des paramètres galactiques, ce qui est particulièrement inhabituel étant donné l'écart d'échelle considérable entre la masse d'un SMBH et celle de sa galaxie hôte (e.g. Magorrian et al. 1998; Kormendy & Ho 2013a; Graham 2016, Ferrarese & Merritt 2000a; Gebhardt et al. 2000a). Ainsi, la mesure de la masse des SMBHs nous permettrait d'enrichir notre compréhension des interactions galactiques. De plus, en effectuant des mesures de masse à des distances plus éloignées, c'est-à-dire dans le passé en tenant compte du temps de propagation de la lumière, nous pourrions approfondir nos connaissances non seulement sur l'état primordial de notre Univers, mais également sur son évolution au cours du temps cosmique.

Il existe plusieurs méthodes de mesure de la masse des SMBHs, généralement regroupées en deux grandes catégories : les méthodes directes et indirectes. Les méthodes directes sont des approches de mesure qui, basées sur des observations, permettent d'obtenir un intervalle précis de la masse. Les méthodes directes discutées dans ce mémoire comprennent les mesures de la cinématique du gaz ionisé, des étoiles, du disque de mégamasers et du gaz moléculaire. Parmi tous ces traceurs cinétiques, le gaz moléculaire offre les mesures les plus précises et exactes. En effet, contrairement aux autres traceurs, il s'agit d'un traceur froid, ce qui implique un mouvement moins chaotique (Davis et al. 2014). De plus, le gaz moléculaire peut être observé par interférométrie, une méthode d'observation qui combine les signaux de plusieurs antennes pour reconstituer une image comme si elle avait été obtenue par une seule antenne de la taille de l'espace entre les antennes. L'interférométrie permet ainsi d'obtenir des mesures plus précises en améliorant la résolution spatiale et en fournissant des informations détaillées sur la cinématique du gaz moléculaire dans la région du SMBH. Cela contribue à

une meilleure compréhension des propriétés et des caractéristiques des SMBHs, ainsi que de leur impact sur l'évolution et l'interaction des galaxies hôtes.

Les méthodes indirectes, quant à elles, relèvent davantage de la prédiction que d'une mesure directe. En utilisant des relations préalablement établies par des mesures directes, il devient possible de prédire rapidement la masse d'un SMBH à partir de l'observation d'un autre paramètre. Par exemple, l'une des corrélations les plus fondamentales et étroitement liées entre un SMBH et sa galaxie hôte est la relation  $M-\sigma$ , qui stipule que la masse d'un SMBH est directement proportionnelle à la dispersion de vitesse stellaire de sa galaxie hôte qui elle est relié à la masse de celle-ci. Ainsi, en mesurant la vitesse de dispersion stellaire d'une galaxie, il est possible d'inférer la masse de son SMBH. Dans ce mémoire, les méthodes indirectes de mesure discutées sont le plan fondamental de l'activité des trous noirs, le profil de brillance de surface et la cartographie par réverbération. Cependant, dans le cadre de ce projet, les méthodes indirectes ne sont pas une option valable, car elles ne fournissent pas de mesure directe et les valeurs utilisées peuvent donc être complètement erronées.

Pour mesurer la masse d'un SMBH en utilisant la cinématique du gaz moléculaire, il est nécessaire d'obtenir des données avec une grande résolution spatiale capable d'observer la sphère d'influence (SOI) du SMBH. La SOI représente la zone où le potentiel gravitationnel du SMBH domine celui de sa galaxie hôte. Étant donné que la masse d'une galaxie est généralement plusieurs ordres de grandeur supérieure à celle d'un SMBH, la SOI ne s'étend que sur quelques parsecs. Les télescopes actuels sont donc capables de mesurer la masse des SMBHs de cette manière uniquement dans l'Univers local, un décalage vers le rouge inférieur à 0.01 (Onishi et al. 2017; North et al. 2019; Davis et al. 2020; Smith et al. 2021a; Lelli et al. 2022). Pour mesurer des masses à des distances plus éloignées, il est nécessaire de trouver un moyen d'améliorer la résolution spatiale de nos données.

Les lentilles gravitationnelles déforment l'image d'une galaxie en arrière-plan lorsqu'elle est observée à travers un objet massif en premier plan. Bien que cette distorsion rende l'objet observé difficile à analyser, elle entraîne également un étirement et un grossissement de l'image. Il devient donc possible de reconstruire une image déformée par lentille gravitationnelle forte avec une résolution spatiale accrue (Meneghetti 2022). Pour ce faire, nous employons un modèle basé sur score (Song et al. 2021), une méthode d'apprentissage automatique capable, tel qu'illustré par Adam et al. 2022, d'inverser l'effet d'une lentille gravitationnelle.

Dans l'article présenté dans ce mémoire, nous abordons la possibilité de mesurer la masse d'un SMBH par la cinématique du gaz moléculaire à un décalage vers le rouge plus élevé, en utilisant l'apprentissage automatique pour tirer avantage du grossissement induit par la distorsion d'une lentille gravitationnelle forte. À cet effet, nous commençons par simuler une observation réaliste du gaz moléculaire capturée par le Grand Réseau d'Antennes Millimétrique/Submillimétrique de l'Atacama (ALMA). Ces données sont générées à partir de la



suite de simulations hydrodynamiques *Feedback In Realistic Environment*, FIRE (Hopkins et al. 2014a). Dans chaque simulation de galaxies, l’effet cinématique d’un SMBH est intégré. Pour ce faire, le gaz moléculaire en orbite autour de celui-ci est considéré comme étant virialisé. Ainsi, grâce au théorème de virial, il est possible de corréler l’énergie potentielle gravitationnelle du SMBH, qui dépend directement de sa masse, à la vitesse du gaz moléculaire. Le flux du gaz moléculaire émis est calculé à partir de cette nouvelle vitesse, de la température, de la densité, de la fraction  $H_2$ , de la distance et de la taille dans le ciel du gaz moléculaire. La propagation du flux à travers le gaz est exprimée en fonction de l’équation de la radiation, en supposant un gradient de vitesse large (LVG). Une fois le flux déterminé, le cube ALMA est généré selon la résolution spatiale et spectrale, qui dépendent du nombre d’antennes, de leur configuration et du temps d’exposition. Enfin, l’effet de lentille gravitationnelle forte est introduit par la rétro-propagation du faisceau lumineux en fonction du profil de masse de l’ellipsoïde isotherme singulière (SIE).

L’exploitation de ces données ALMA simulées est testée sur un problème de régression directe. Nous entraînons un réseau de neurones à convolution (CNN) pour apprendre à prédire avec précision, exactitude et rapidité la masse d’un SMBH à partir des données simulées, sans prendre en compte l’effet de la lentille. Le réseau prédit la masse du SMBH ainsi que son incertitude en supposant une distribution a posteriori gaussienne. Les résultats sont convaincants : plus la masse du SMBH est élevée, plus la prédiction du réseau est précise et exacte. Tout comme avec les méthodes conventionnelles, le réseau est uniquement capable de prédire la masse du SMBH tant que la résolution spatiale des données permet de résoudre la SOI. De plus, les cartes de saillance du réseau confirment que celui-ci utilise l’information dans la SOI pour prédire la masse du SMBH. Nos travaux futurs se concentreront sur l’exploration de données simulées d’ALMA affectées par l’effet de lentilles gravitationnelles fortes, dans le but d’étudier la faisabilité de mesurer les masses des SMBHs à des décalage vers le rouge plus élevés. Notre approche consistera à reconstruire la galaxie située en arrière-plan en tirant profit de l’amplification produite par l’effet de lentille, ce qui nous permettrait de surmonter les limitations liées à la résolution spatiale. Le cube ainsi reconstruit pourra alors être directement introduit dans le réseau que nous avons préalablement entraîné, comme décrit dans l’article associé à ce mémoire, pour mesurer la masse du SMBH.

## Bibliographie

- Abbott, R., Abbott, T. D., Acernese, F., et al. 2022, *Astronomy & Astrophysics*, 659, A84
- Adam, A., Coogan, A., Malkin, N., et al. 2022, arXiv e-prints, arXiv:2211.03812
- Alatalo, K., Davis, T. A., Bureau, M., et al. 2013, , 432, 1796
- and A. Amorim, Bourdarot, G., Brandner, W., et al. 2022, *Astronomy & Astrophysics*, 669, A14
- Anglés-Alcázar, D., Faucher-Giguère, C.-A., Kereš, D., et al. 2017a, , 470, 4698
- Anglés-Alcázar, D., Faucher-Giguère, C.-A., Quataert, E., et al. 2017b, , 472, L109
- Atek, H., Shuntov, M., Furtak, L. J., et al. 2023, , 519, 1201
- Barnes, J. & Hut, P. 1986, , 324, 446
- Barth, A. J., Boizelle, B., Darling, J. K., et al. 2016a, in *American Astronomical Society Meeting Abstracts*, Vol. 228, *American Astronomical Society Meeting Abstracts #228*, 103.05
- Barth, A. J., Boizelle, B., Darling, J. K., et al. 2016b, in *American Astronomical Society Meeting Abstracts*, Vol. 228, *American Astronomical Society Meeting Abstracts #228*, 103.05
- Barth, A. J., Boizelle, B. D., Darling, J., et al. 2016c, , 822, L28
- Barth, A. J., Sarzi, M., Rix, H.-W., et al. 2001, , 555, 685
- Bayes, T. 1763, *Philosophical Transactions of the Royal Society of London Series I*, 53, 269
- Bayliss, M. B., McDonald, M., Sharon, K., et al. 2020, *Nature Astronomy*, 4, 159
- Begelman, M. C., Blandford, R. D., & Rees, M. J. 1980, , 287, 307
- Belokurov, V., Sanders, J. L., Fattahi, A., et al. 2020, , 494, 3880
- Bengio, Y., Lee, D.-H., Bornschein, J., Mesnard, T., & Lin, Z. 2015, arXiv e-prints, arXiv:1502.04156
- Bengio, Y., Léonard, N., & Courville, A. 2013, arXiv e-prints, arXiv:1308.3432
- Benjamin, R. A., Churchwell, E., Babler, B. L., et al. 2005, *The Astrophysical Journal*, 630, L149
- Bentz, M. C., Denney, K. D., Grier, C. J., et al. 2013, , 767, 149
- Bernardi, M., Sheth, R. K., Tundo, E., & Hyde, J. B. 2007, , 660, 267
- Binney, J. & Tremaine, S. 2008, *Galactic Dynamics : Second Edition*
- Birrer, S. & Amara, A. 2018, *Lenstronomy : multi-purpose gravitational lens modelling software package*, arXiv:1803.09746 [astro-ph.CO]
- Birrer, S., Shajib, A. J., Gilman, D., et al. 2021, *Journal of Open Source Software*, 6, 3283
- Boizelle, B. D., Barth, A. J., Darling, J., et al. 2017, , 845, 170
- Boizelle, B. D., Barth, A. J., Walsh, J. L., et al. 2019, , 881, 10
- Boltzmann, L. 1894, *Annalen der Physik*, 289, 959

Bom, C., Poh, J., Nord, B., Blanco-Valentin, M., & Dias, L. 2019, arXiv e-prints, arXiv:1911.06341

Bower, R. G., Benson, A. J., Malbon, R., et al. 2006, , 370, 645

Brockamp, M., Baumgardt, H., & Kroupa, P. 2011, , 418, 1308

Bullock, J. S., Kolatt, T. S., Sigad, Y., et al. 2001, , 321, 559

Caldwell, J. A. R. & Ostriker, J. P. 1981, , 251, 61

CASA Team, Bean, B., Bhatnagar, S., et al. 2022, , 134, 114501

Çatmabacak, O., Feldmann, R., Anglés-Alcázar, D., et al. 2022, , 511, 506

Charnock, T., Perreault-Levasseur, L., & Lanusse, F. 2020, arXiv e-prints, arXiv:2006.01490

Chen, Y.-C., Ho, S., Freeman, P. E., Genovese, C. R., & Wasserman, L. 2015, , 454, 1140

Cheng, T.-Y., Conselice, C. J., Aragón-Salamanca, A., et al. 2021, , 507, 4425

Clevert, D.-A., Unterthiner, T., & Hochreiter, S. 2015, arXiv e-prints, arXiv:1511.07289

Combes, F., Young, L. M., & Bureau, M. 2007, , 377, 1795

Combes, F., García-Burillo, S., Audibert, A., et al. 2019, , 623, A79

Crain, R. A., Schaye, J., Bower, R. G., et al. 2015, , 450, 1937

Davé, R., Anglés-Alcázar, D., Narayanan, D., et al. 2019, , 486, 2827

Davis, B. L., Graham, A. W., & Cameron, E. 2018a, , 869, 113

Davis, B. L., Graham, A. W., & Cameron, E. 2019, , 873, 85

Davis, T. A., Bureau, M., Cappellari, M., Sarzi, M., & Blitz, L. 2013, , 494, 328

Davis, T. A., Bureau, M., Onishi, K., et al. 2017, , 468, 4675

Davis, T. A., Alatalo, K., Sarzi, M., et al. 2011, , 417, 882

Davis, T. A., Young, L. M., Crocker, A. F., et al. 2014, , 444, 3427

Davis, T. A., Bureau, M., Onishi, K., et al. 2018b, , 473, 3818

Davis, T. A., Nguyen, D. D., Seth, A. C., et al. 2020, , 496, 4061

Dehnen, W. & Read, J. I. 2011, *European Physical Journal Plus*, 126, 55

Deserno, M. & Holm, C. 1998, , 109, 7678

DeVries, T. & Taylor, G. W. 2017, arXiv e-prints, arXiv:1708.04552

Di Matteo, T., Springel, V., & Hernquist, L. 2005, , 433, 604

Dieleman, S., Willett, K. W., & Dambre, J. 2015, , 450, 1441

Domínguez Sánchez, H., Huertas-Company, M., Bernardi, M., Tuccillo, D., & Fischer, J. L. 2018, , 476, 3661

Draine, B. T., Aniano, G., Krause, O., et al. 2014, , 780, 172

Dressler, A. 1980, , 236, 351

Dressler, A. 1989, *Annals of the New York Academy of Sciences*, 571, 308

Dubinski, J., Kim, J., Park, C., & Humble, R. 2004, , 9, 111

Dunlop, J. S., McLure, R. J., Kukula, M. J., et al. 2003, , 340, 1095

Einstein, A. 1905, *Annalen der Physik*, 322, 549

Einstein, A. 1936, *Science*, 84, 506

Einstein, A., Infeld, L., & Hoffmann, B. 1938, *Annals of Mathematics*, 39, 65

Eriksen, M., Alarcon, A., Cabayol, L., et al. 2020, , 497, 4565

Euclid Collaboration, Scaramella, R., Amiaux, J., et al. 2022, , 662, A112

Euler, H. 1936, *Annalen der Physik*, 418, 398

Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al. 2019, , 875, L1

Fabian, A. C. & Lasenby, A. N. 2015, in *General Relativity : The Most Beautiful of Theories*, ed. C. Rovelli, Vol. 28, 7

Falcke, H., K rding, E., & Markoff, S. 2004, , 414, 895

Farrah, D., Croker, K. S., Zevin, M., et al. 2023, , 944, L31

Feldmann, R., Hopkins, P. F., Quataert, E., Faucher-Gigu re, C.-A., & Kere , D. 2016, , 458, L14

Feldmann, R., Quataert, E., Hopkins, P. F., Faucher-Gigu re, C.-A., & Kere , D. 2017, , 470, 1050

Ferrarese, L. & Merritt, D. 2000a, , 539, L9

Ferrarese, L. & Merritt, D. 2000b, , 539, L9

Ford, H. C., Harms, R. J., Tsvetanov, Z. I., et al. 1994, , 435, L27

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, , 125, 306

Friedmann, A. 1922, *Zeitschrift fur Physik*, 10, 377

Gebhardt, K., Bender, R., Bower, G., et al. 2000a, , 539, L13

Gebhardt, K., Kormendy, J., Ho, L. C., et al. 2000b, , 543, L5

Ghez, A. M., Salim, S., Weinberg, N. N., et al. 2008, , 689, 1044

Gingold, R. A. & Monaghan, J. J. 1977, , 181, 375

Goodfellow, I. J., Bengio, Y., & Courville, A. 2016, *Deep Learning* (Cambridge, MA, USA : MIT Press), <http://www.deeplearningbook.org>

Goodfellow, I. J., Courville, A., & Bengio, Y. 2013, arXiv e-prints, arXiv:1301.3568

Gow, A. D., Byrnes, C. T., & Hall, A. 2022, , 105, 023503

Graham, A. W. 2016, in *Astrophysics and Space Science Library*, Vol. 418, *Galactic Bulges*, ed. E. Laurikainen, R. Peletier, & D. Gadotti, 263

Graham, A. W. & Scott, N. 2013, , 764, 151

Greene, J. E. & Ho, L. C. 2005, , 630, 122

Greene, J. E., Peng, C. Y., Kim, M., et al. 2010, , 721, 26

Grudi , M. Y., Hafen, Z., Rodriguez, C. L., et al. 2023, , 519, 1366

Gultekin, K. 2009, arXiv e-prints, arXiv :0912.3898

Harnois-D raps, J., Pen, U.-L., Iliev, I. T., et al. 2013, , 436, 540

Hausen, R. & Robertson, B. 2022, arXiv e-prints, arXiv:2201.04714

He, K., Zhang, X., Ren, S., & Sun, J. 2015, arXiv e-prints, arXiv:1512.03385

Heckman, T. M. & Best, P. N. 2014, , 52, 589

Hernquist, L. 1987, , 64, 715

Hernquist, L. & Katz, N. 1989, , 70, 419

Hezaveh, Y. D. 2014, , 791, L41

Hezaveh, Y. D., Perreault Levasseur, L., & Marshall, P. J. 2017a, , 548, 555

Hezaveh, Y. D., Perreault Levasseur, L., & Marshall, P. J. 2017b, , 548, 555

Hezaveh, Y. D., Dalal, N., Marrone, D. P., et al. 2016, , 823, 37

Hlavacek-Larrondo, J., Fabian, A. C., Edge, A. C., et al. 2012a, , 421, 1360

Hlavacek-Larrondo, J., Fabian, A. C., Edge, A. C., & Hogan, M. T. 2012b, , 424, 224

Hoekstra, H., Yee, H. K. C., & Gladders, M. D. 2004, , 606, 67

Hopkins, P. F. 2015, , 450, 53

Hopkins, P. F. 2016, , 462, 576

Hopkins, P. F., Kereš, D., Oñorbe, J., et al. 2014a, , 445, 581

Hopkins, P. F., Kereš, D., Oñorbe, J., et al. 2014b, , 445, 581

Hopkins, P. F., Wetzel, A., Kereš, D., et al. 2018, , 480, 800

Hoyle, B. 2016, *Astronomy and Computing*, 16, 34

Hubble, E. 1929, *Proceedings of the National Academy of Science*, 15, 168

Hubble, E. P. 1936, *Realm of the Nebulae*

Huber, S., Suyu, S. H., Ghoshdastidar, D., et al. 2022, , 658, A157

Huertas-Company, M., Gravet, R., Cabrera-Vives, G., et al. 2015, , 221, 8

Ioffe, S. & Szegedy, C. 2015, arXiv e-prints, arXiv:1502.03167

Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, , 873, 111

Kaspi, S., Smith, P. S., Netzer, H., et al. 2000, , 533, 631

Kereš, D., Katz, N., Weinberg, D. H., & Davé, R. 2005, , 363, 2

Kerr, R. P. 1963, , 11, 237

Kingma, D. P. & Ba, J. 2014, arXiv e-prints, arXiv:1412.6980

Klypin, A. & Holtzman, J. 1997, arXiv e-prints, astro

Kobayashi, C. 2004, , 347, 740

Körding, E., Falcke, H., & Corbel, S. 2006, , 456, 439

Kormendy, J. & Ho, L. C. 2013a, , 51, 511

Kormendy, J. & Ho, L. C. 2013b, , 51, 511

Kormendy, J. & Richstone, D. 1995, , 33, 581

Kuo, C. Y., Braatz, J. A., Condon, J. J., et al. 2011, , 727, 20

Läscher, R., Greene, J. E., Seth, A., et al. 2016, , 825, 3

Latif, M. A. & Ferrara, A. 2016, , 33, e051

Lauer, T. R., Tremaine, S., Richstone, D., & Faber, S. M. 2007a, , 670, 249

Lauer, T. R., Faber, S. M., Richstone, D., et al. 2007b, , 662, 808

LeCun, Y., Bengio, Y., & Hinton, G. 2015, , 521, 436

Legin, R., Adam, A., Hezaveh, Y., & Perreault Levasseur, L. 2023, arXiv e-prints, arXiv:2302.03046

Lei Ba, J., Kiros, J. R., & Hinton, G. E. 2016, arXiv e-prints, arXiv:1607.06450  
 Lelli, F., Davis, T. A., Bureau, M., et al. 2022, , 516, 4066  
 Lemon, C., Anguita, T., Auger-Williams, M. W., et al. 2023, , 520, 3305  
 Leroy, A. K., Rosolowsky, E., Usero, A., et al. 2022, , 927, 149  
 Li, R., Napolitano, N. R., Roy, N., et al. 2022, , 929, 152  
 Li, X. & Mandelbaum, R. 2023, , 521, 4904  
 Lin, M., Chen, Q., & Yan, S. 2013, arXiv e-prints, arXiv:1312.4400  
 Loveday, J. 1996, , 278, 1025  
 Ma, X., Hopkins, P. F., Faucher-Giguère, C.-A., et al. 2016, , 456, 2140  
 MacFarland, T., Couchman, H. M. P., Pearce, F. R., & Pichlmeier, J. 1998, , 3, 687  
 Madireddy, S., Ramachandra, N., Li, N., et al. 2019, arXiv e-prints, arXiv:1911.03867  
 Magorrian, J., Tremaine, S., Richstone, D., et al. 1998, , 115, 2285  
 Marconi, A., Pastorini, G., Pacini, F., et al. 2006, , 448, 921  
 Marconi, M. L. 2003, , 166, 410  
 Markoff, S. & Event Horizon Telescope Collaboration. 2022, in American Astronomical Society Meeting Abstracts, Vol. 54, American Astronomical Society Meeting Abstracts, 211.01  
 Mathieu, M., Couprie, C., & LeCun, Y. 2015, arXiv e-prints, arXiv:1511.05440  
 McConnell, N. J., Ma, C., Graham, J. R., et al. 2011a, in American Astronomical Society Meeting Abstracts, Vol. 217, American Astronomical Society Meeting Abstracts #217, 422.06  
 McConnell, N. J. & Ma, C.-P. 2013, , 764, 184  
 McConnell, N. J., Ma, C.-P., Gebhardt, K., et al. 2011b, , 480, 215  
 McConnell, N. J., Ma, C.-P., Murphy, J. D., et al. 2012, , 756, 179  
 McLure, R. J. & Jarvis, M. J. 2002, Monthly Notices of the Royal Astronomical Society, 337, 109  
 Mehrgan, K., Thomas, J., Saglia, R., et al. 2019, , 887, 195  
 Meneghetti, M. 2022, Introduction to Gravitational Lensing : With Python Examples  
 Merloni, A., Heinz, S., & di Matteo, T. 2003, , 345, 1057  
 Mezcuca, M., Civano, F., Marchesi, S., et al. 2018a, , 478, 2576  
 Mezcuca, M., Hlavacek-Larrondo, J., Lucey, J. R., et al. 2018b, , 474, 1342  
 Mihalas, D. & Routly, P. M. 1968, Galactic astronomy  
 Miyoshi, M., Moran, J., Herrnstein, J., et al. 1995, , 373, 127  
 Mlotshwa, T., van Deventer, H., & Bosman, A. S. 2023, Cauchy Loss Function : Robustness Under Gaussian and Cauchy Noise, arXiv:2302.07238 [cs.LG]  
 Moriyama, K., Mineshige, S., Honma, M., & Akiyama, K. 2019, The Astrophysical Journal, 887, 227  
 Morningstar, W. R., Perreault Levasseur, L., Hezaveh, Y. D., et al. 2019, , 883, 14

- Nair, D. S., Hochgeschwender, N., & Olivares-Mendez, M. A. 2022, arXiv e-prints, arXiv:2202.03870
- Navarro, J. F., Eke, V. R., & Frenk, C. S. 1996, , 283, L72
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, , 490, 493
- Nguyen, D. D., den Brok, M., Seth, A. C., et al. 2020, , 892, 68
- North, E. V., Davis, T. A., Bureau, M., et al. 2019, , 490, 319
- Onishi, K., Iguchi, S., Davis, T. A., et al. 2017, , 468, 4663
- Onken, C. A., Ferrarese, L., Merritt, D., et al. 2004, , 615, 645
- Pasquet, J., Bertin, E., Treyer, M., Arnouts, S., & Fouchez, D. 2019, , 621, A26
- Pearson, J., Li, N., & Dye, S. 2019, , 488, 991
- Pensabene, A., Carniani, S., Perna, M., et al. 2020, , 637, A84
- Perreault Levasseur, L., Hezaveh, Y. D., & Wechsler, R. H. 2017, , 850, L7
- Peterson, B. M. 1993, , 105, 247
- Peterson, B. M., Ferrarese, L., Gilbert, K. M., et al. 2004, , 613, 682
- Pillepich, A., Springel, V., Nelson, D., et al. 2018, , 473, 4077
- Piotrowska, J. M., Bluck, A. F. L., Maiolino, R., & Peng, Y. 2022, , 512, 1052
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, , 594, A13
- Plotkin, R. M., Markoff, S., Kelly, B. C., K rding, E., & Anderson, S. F. 2012, , 419, 267
- Poisson. 1824, *Annalen der Physik*, 77, 301
- Price, D. J. 2012, , 420, L33
- Recht, B., Roelofs, R., Schmidt, L., & Shankar, V. 2018, arXiv e-prints, arXiv:1806.00451
- Rein, H. & Liu, S. F. 2012, , 537, A128
- Remy, B., Lanusse, F., Jeffrey, N., et al. 2023, , 672, A51
- Rhea, C., Hlavacek-Larrondo, J., Rousseau-Nepton, L., & Prunet, S. 2021, *Research Notes of the American Astronomical Society*, 5, 276
- Riffel, R. A. & Storchi-Bergmann, T. 2011, , 411, 469
- Robert, C. & Casella, G. 2011, *Statistical Science*, 26
- Ruffa, I., Davis, T. A., Cappellari, M., et al. 2023, , 522, 6170
- Rusli, S. P., Erwin, P., Saglia, R. P., et al. 2013, , 146, 160
- Sahu, N., Graham, A. W., & Davis, B. L. 2019, , 876, 155
- Sani, E., Marconi, A., Hunt, L. K., & Risaliti, G. 2011, , 413, 1479
- Schawinski, K., Urry, C. M., Virani, S., et al. 2010, , 711, 284
- Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, , 446, 521
- Schmidhuber, J. 2014, arXiv e-prints, arXiv:1404.7828
- Schnee, S. L., Brogan, C., Espada, D., et al. 2014, in *SPIE Proceedings*, ed. A. B. Peck, C. R. Benn, & R. L. Seaman (SPIE)
- Schwarzschild, K. 1916, *Abh. Konigl. Preuss. Akad. Wissenschaften Jahre 1906,92*, Berlin,1907, 1916, 189

- Sérsic, J. L. 1963, *Boletin de la Asociacion Argentina de Astronomia La Plata Argentina*, 6, 41
- Sexton, R. O., Canalizo, G., Hiner, K. D., et al. 2019, , 878, 101
- Shankar, F., Bernardi, M., Sheth, R. K., et al. 2016, in *Active Galactic Nuclei: What's in a Name?*, 65
- Shrivastava, A., Sukthankar, R., Malik, J., & Gupta, A. 2016, arXiv e-prints, arXiv:1612.06851
- Silk, J. 1998, in *The Hubble Deep Field*, ed. M. Livio, S. M. Fall, & P. Madau, 194
- Simonyan, K. & Zisserman, A. 2014, arXiv e-prints, arXiv:1409.1556
- Smith, M. D., Bureau, M., Davis, T. A., et al. 2019, , 485, 4359
- Smith, M. D., Bureau, M., Davis, T. A., et al. 2021a, , 500, 1933
- Smith, M. D., Bureau, M., Davis, T. A., et al. 2021b, , 503, 5984
- Sohl-Dickstein, J., Weiss, E., Maheswaranathan, N., & Ganguli, S. 2015, in *Proceedings of Machine Learning Research*, Vol. 37, *Proceedings of the 32nd International Conference on Machine Learning*, ed. F. Bach & D. Blei (Lille, France : PMLR), 2256
- Song, Y. & Ermon, S. 2019, arXiv e-prints, arXiv:1907.05600
- Song, Y., Sohl-Dickstein, J., Kingma, D. P., et al. 2021, *Score-Based Generative Modeling through Stochastic Differential Equations*, arXiv:2011.13456 [cs.LG]
- Sonnenfeld, A. 2022, *Astronomy & Astrophysics*, 659, A141
- Springel, V. 2005, , 364, 1105
- Springel, V. 2010, , 48, 391
- Springel, V. & Hernquist, L. 2003, , 339, 289
- Springenberg, J. T., Dosovitskiy, A., Brox, T., & Riedmiller, M. 2014, arXiv e-prints, arXiv:1412.6806
- Stinson, G., Seth, A., Katz, N., et al. 2006, , 373, 1074
- Sun, R. 2019, arXiv e-prints, arXiv:1912.08957
- Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. 2016, arXiv e-prints, arXiv:1602.07261
- Szegedy, C., Liu, W., Jia, Y., et al. 2014, arXiv e-prints, arXiv:1409.4842
- Tadaki, K.-i., Iye, M., Fukumoto, H., et al. 2020, , 496, 4276
- Theuns, T., Leonard, A., Efstathiou, G., Pearce, F. R., & Thomas, P. A. 1998, , 301, 478
- Tremaine, A., Wang, X. J., Babzien, M., et al. 2002, , 66, 036503
- Tuccillo, D., Huertas-Company, M., Decencière, E., et al. 2018, , 475, 894
- Valiev, M., Bylaska, E. J., Govind, N., et al. 2010, *Computer Physics Communications*, 181, 1477
- van den Bosch, R. C. E. 2016, , 831, 134
- van den Bosch, R. C. E. & de Zeeuw, P. T. 2010, , 401, 1770
- van der Marel, R. P. & van den Bosch, F. C. 1998, , 116, 2220
- Vestergaard, M. & Peterson, B. M. 2006, , 641, 689



Villanueva-Domingo, P., Villaescusa-Navarro, F., Anglés-Alcázar, D., et al. 2022, , 935, 30  
Vogelsberger, M., Genel, S., Springel, V., et al. 2014, , 444, 1518  
Walsh, J. L., Barth, A. J., Ho, L. C., & Sarzi, M. 2013, , 770, 86  
Wetzel, A., Hayward, C. C., Sanderson, R. E., et al. 2023, , 265, 44  
Willett, K. W., Lintott, C. J., Bamford, S. P., et al. 2013, , 435, 2835  
Y Mueller, B., Roth, T., Cinchetti, M., Aeschlimann, M., & Rethfeld, B. 2011, *New Journal of Physics*, 13, 123010  
You, Y., Gitman, I., & Ginsburg, B. 2017, arXiv e-prints, arXiv:1708.03888  
Young, L. M., Bureau, M., Davis, T. A., et al. 2011, , 414, 940  
Zhang, C., Vinyals, O., Munos, R., & Bengio, S. 2018, arXiv e-prints, arXiv:1804.06893  
Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., & Torralba, A. 2015, arXiv e-prints, arXiv:1512.04150  
Zou, D., Cao, Y., Zhou, D., & Gu, Q. 2018, arXiv e-prints, arXiv:1811.08888