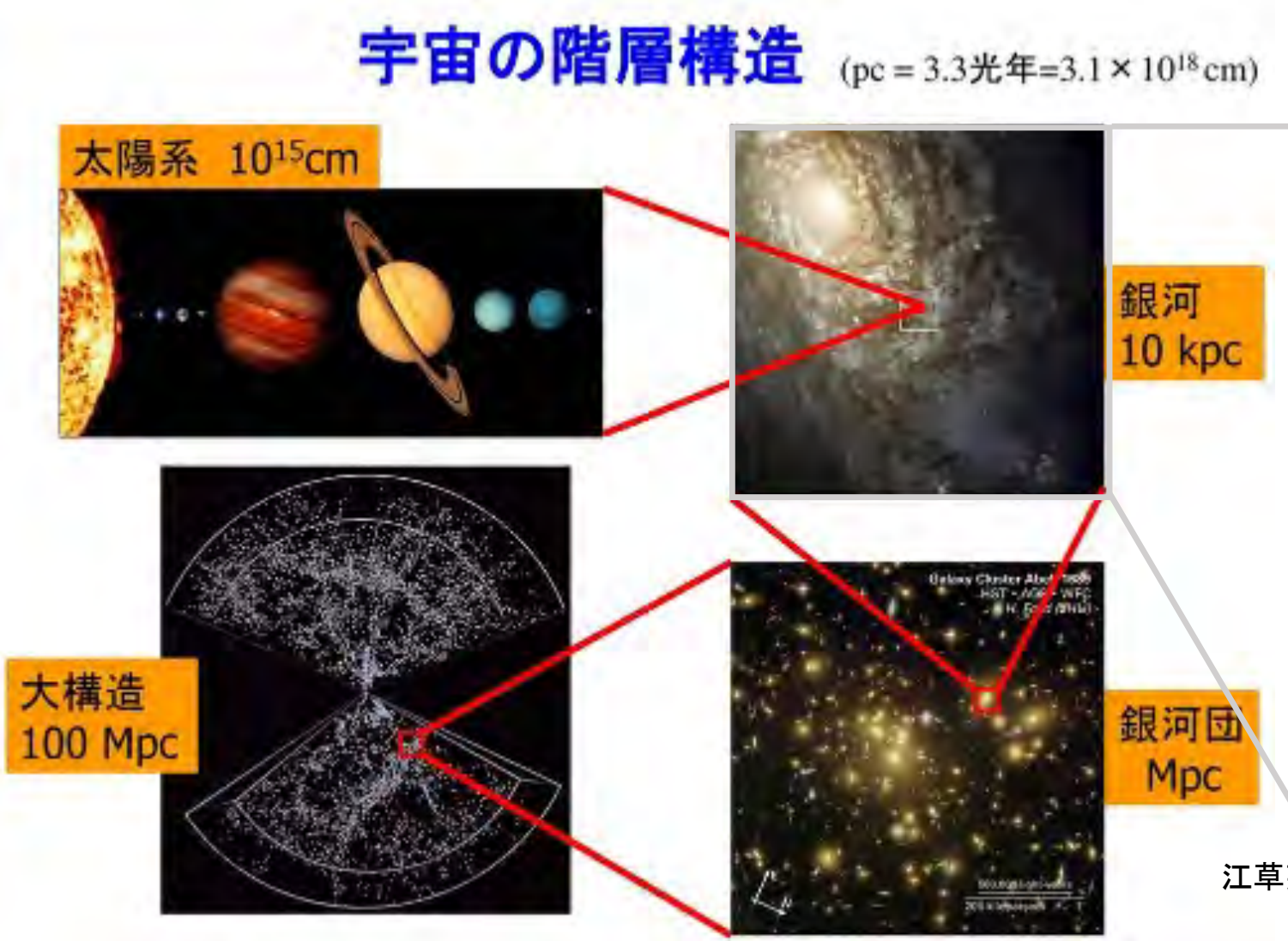
「機械学習」と「XAI」で挑む 現象分離手法の確立

D2 大久保宏真（筑波大学 数理物質系,宇宙観測研究室）

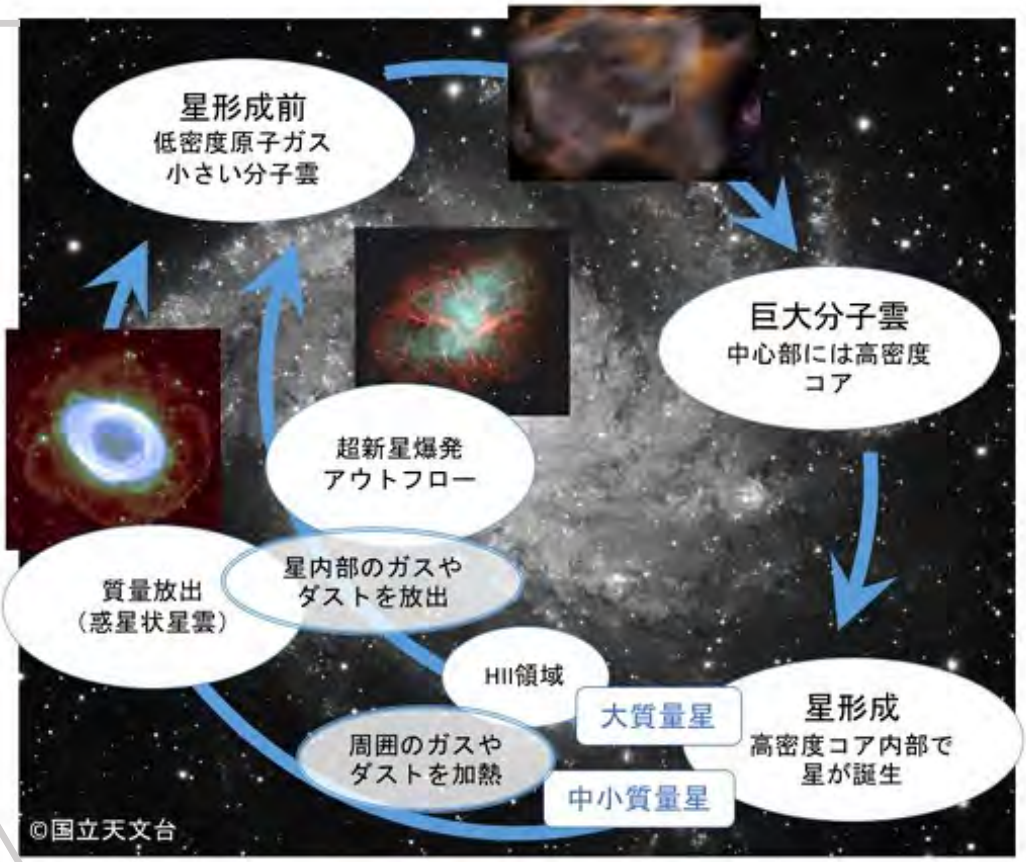
共同研究者：竹内努（名古屋大学、統数研）、赤穂昭太郎（産総研、統数研）、
五十嵐康彦（筑波大学 システム情報系）、久野成夫（筑波大学 数理物質系）、斉藤俊貴（静岡大学）、
高野秀路（日本大学）、谷口暁星（北見工業大学）、中島拓（公立諏訪東京理科大学）、原田ななせ（国立天文台）

Okubo et al. (2025)
Okubo et al. in prep
Okubo et al. in prep

宇宙の階層構造と銀河



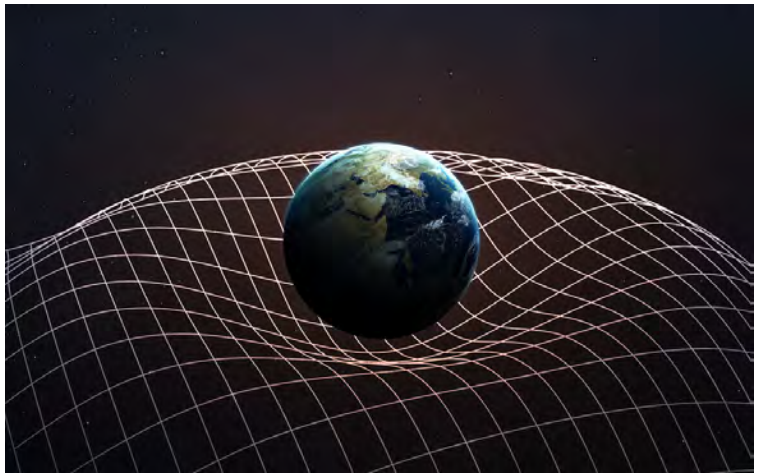
北山哲 – 大規模構造の起源



江草芙実 – SPICAで見る近傍銀河と銀河系, 天文月報 第113巻 第12号

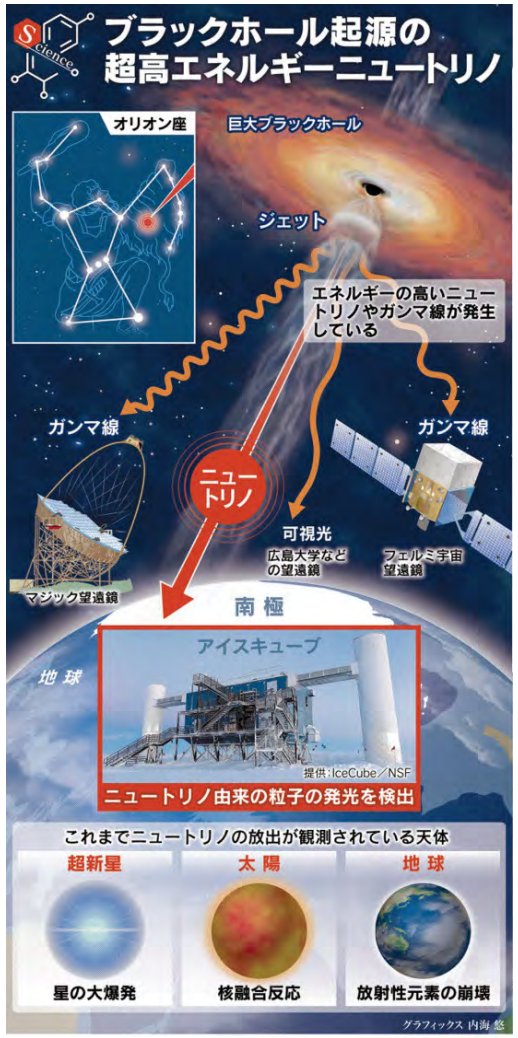
どうやって調べるの？

・重力波観測



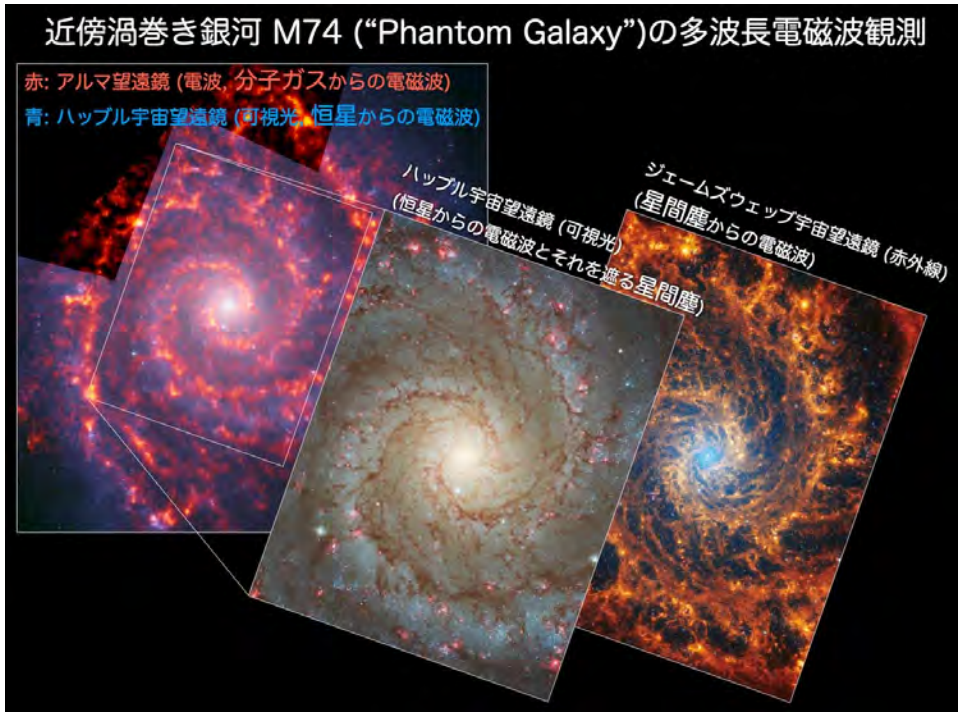
重力波の2回目の観測に成功 – ロシア・ビヨンド

・ニュートリノ観測



「ニュートリノ」放出する新天体　南極の施設で観測 – 日本経済新聞

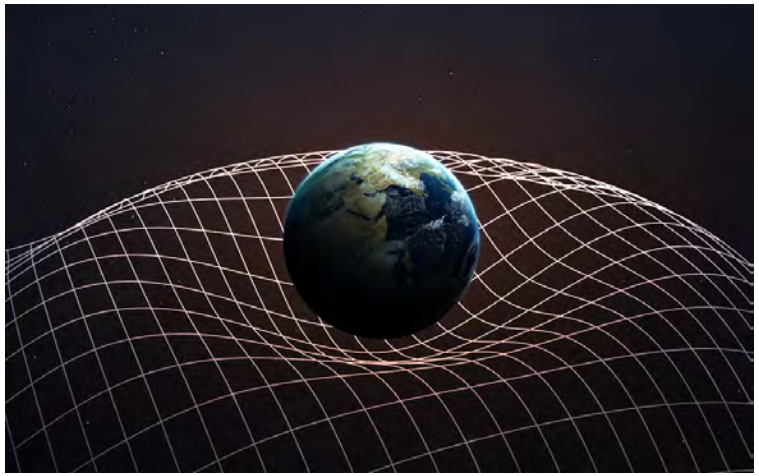
・電磁波観測



斉藤俊貴 – 静岡大学 グローバル共創科学部

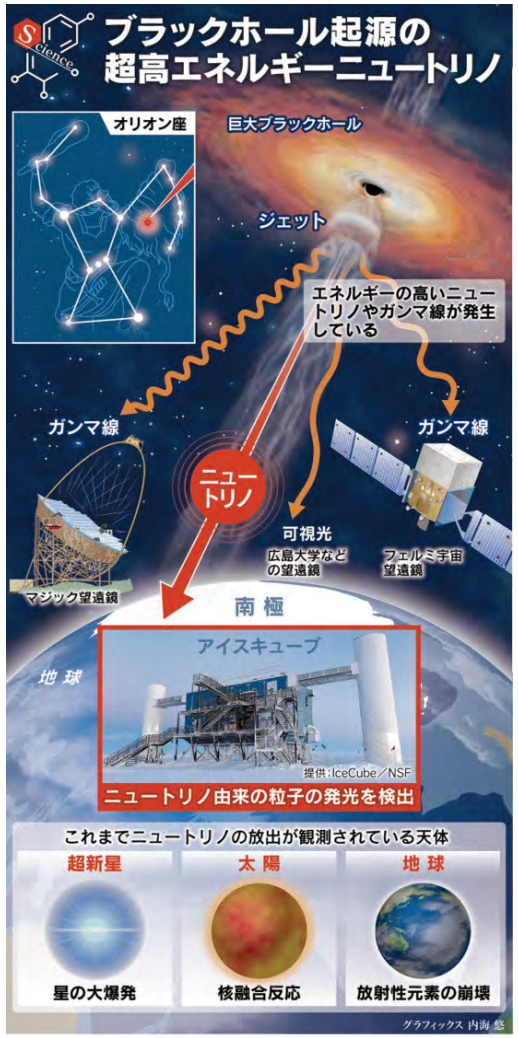
どうやって調べるの？

・重力波観測



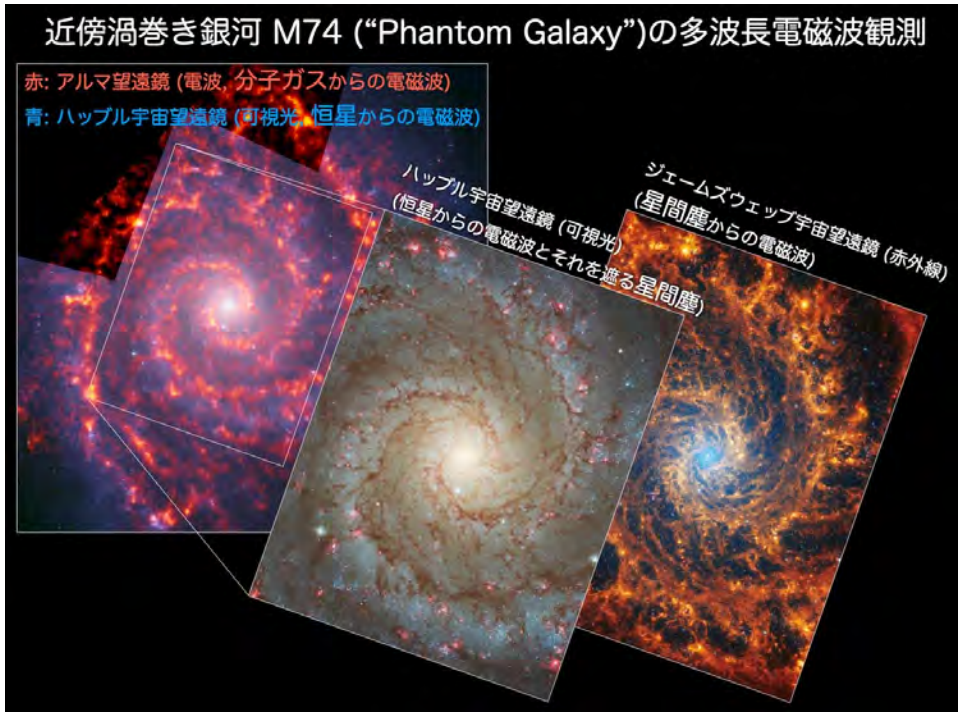
重力波の2回目の観測に成功 – ロシア・ビヨンド

・ニュートリノ観測



「ニュートリノ」放出する新天体　南極の施設で観測 – 日本経済新聞

・電磁波観測

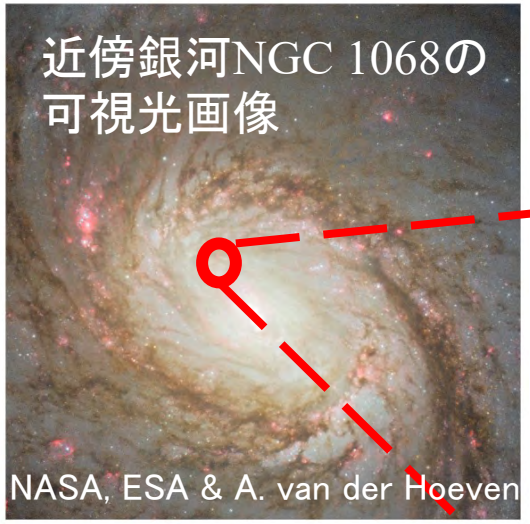


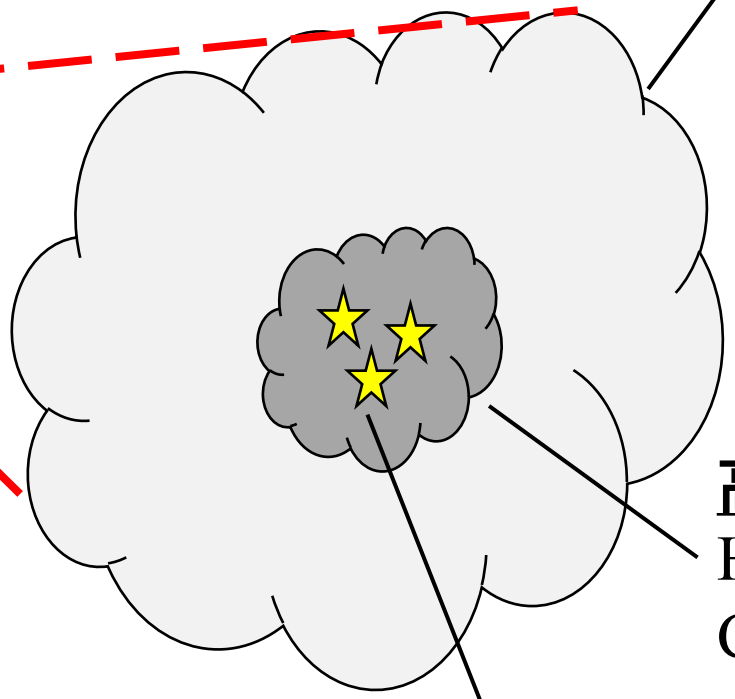
斉藤俊貴 – 静岡大学 グローバル共創科学部

→ **電波に注目！**

- ダスト減光を受けにくい。
- 分子が放射する光（分子輝線）をたくさん観測できる。
- etc

分子輝線の特徴





低密度ガスをトレース（10^2–10^3 /cm^{-3}）

^{12}CO (J=1-0), ^{13}CO (J=1-0)

高密度ガスをトレース（10^4–10^5 /cm^{-3}）

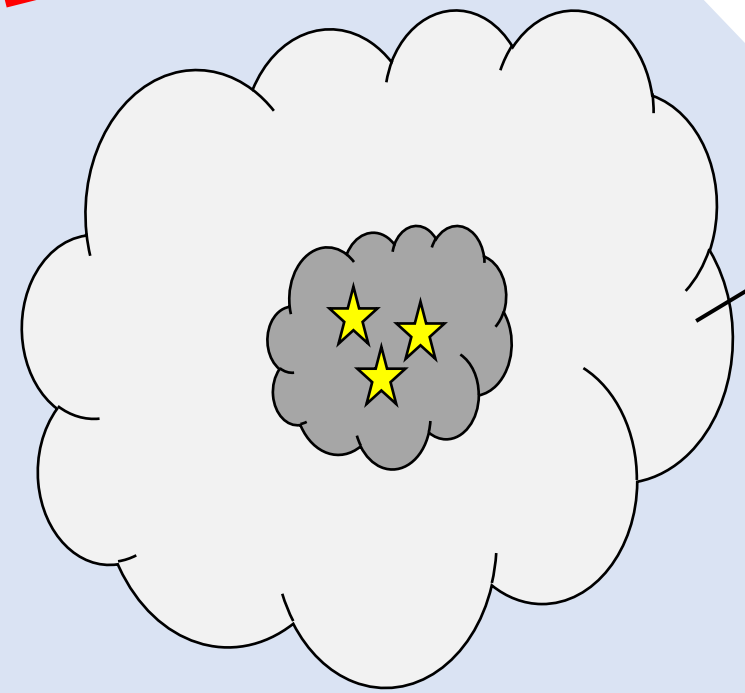
HCN (J=1-0), CN (N=1-0),
C_2H (N=1-0), N_2H^+ (J=1-0)

光解離領域（PDR）

COが破壊

→ CN, C_2Hが生成

分子輝線の特徴



COが破壊
→ CN, C_2Hが生成

たくさんの分子輝線を見ること

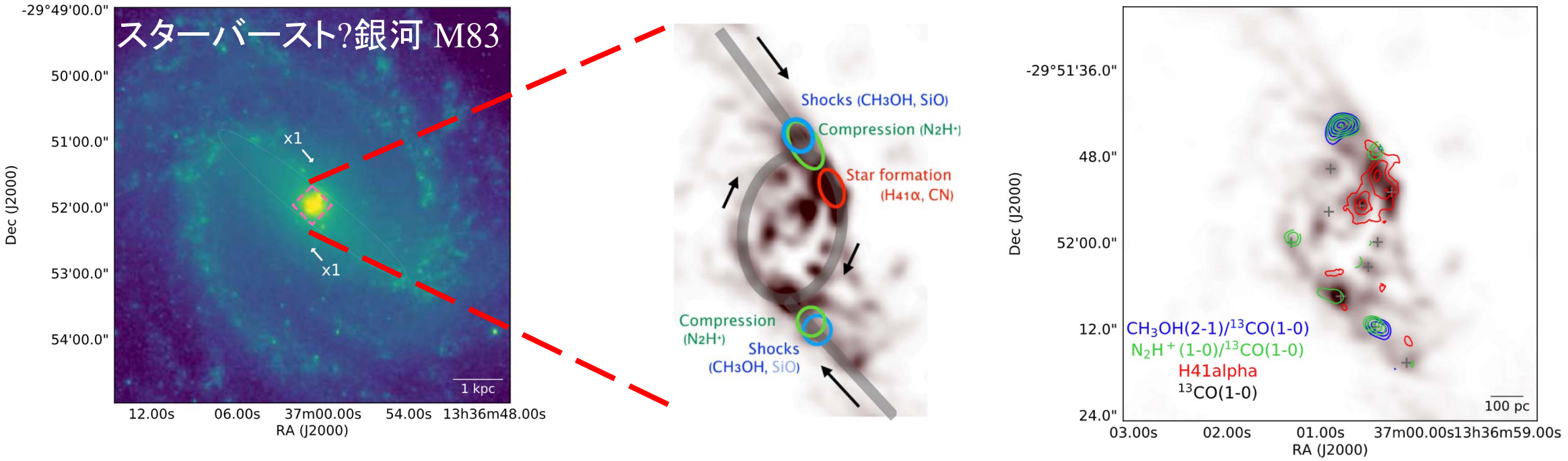
=

銀河内部で発生している現象を知る
手掛かりを得ること

⇩

多輝線解析は重要！

多輝線解析



(Harada et al. 2019)

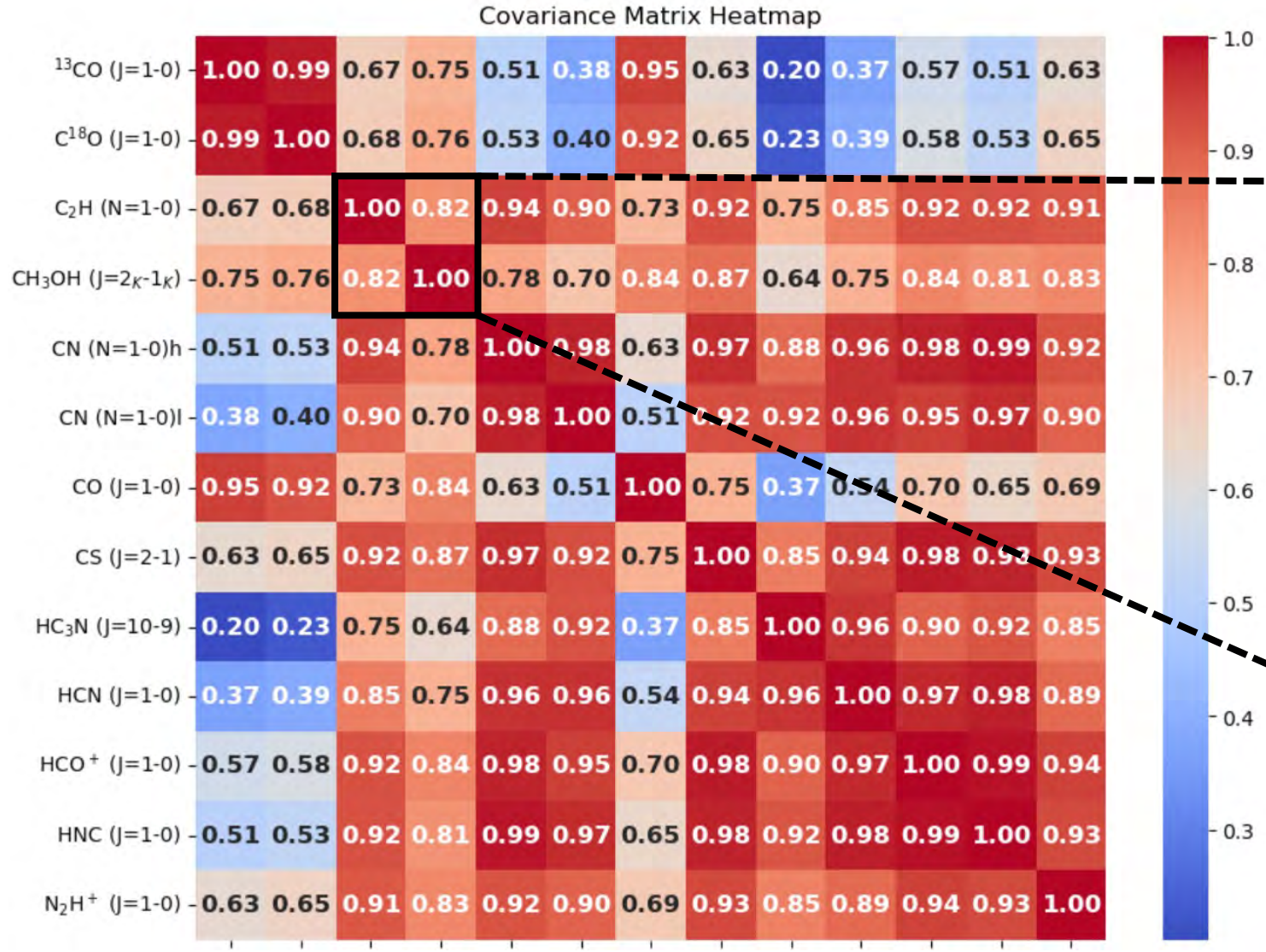
従来の多輝線解析

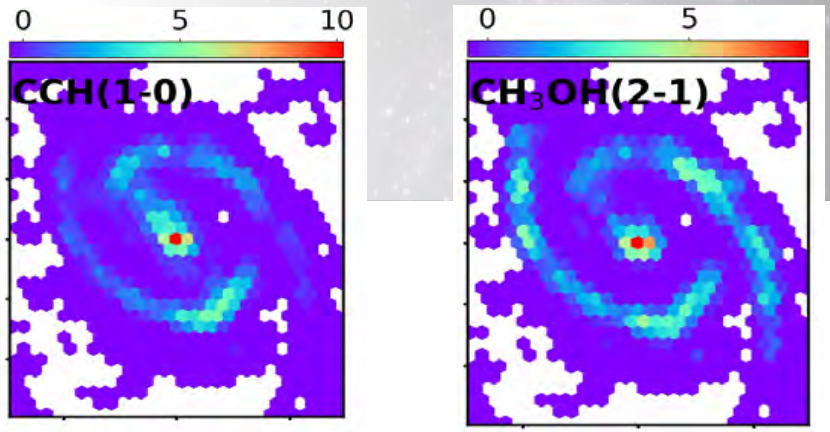
- 輝線比により、物理量を間接的に反映した情報を可視化
- 輝線間の関係性の把握は困難

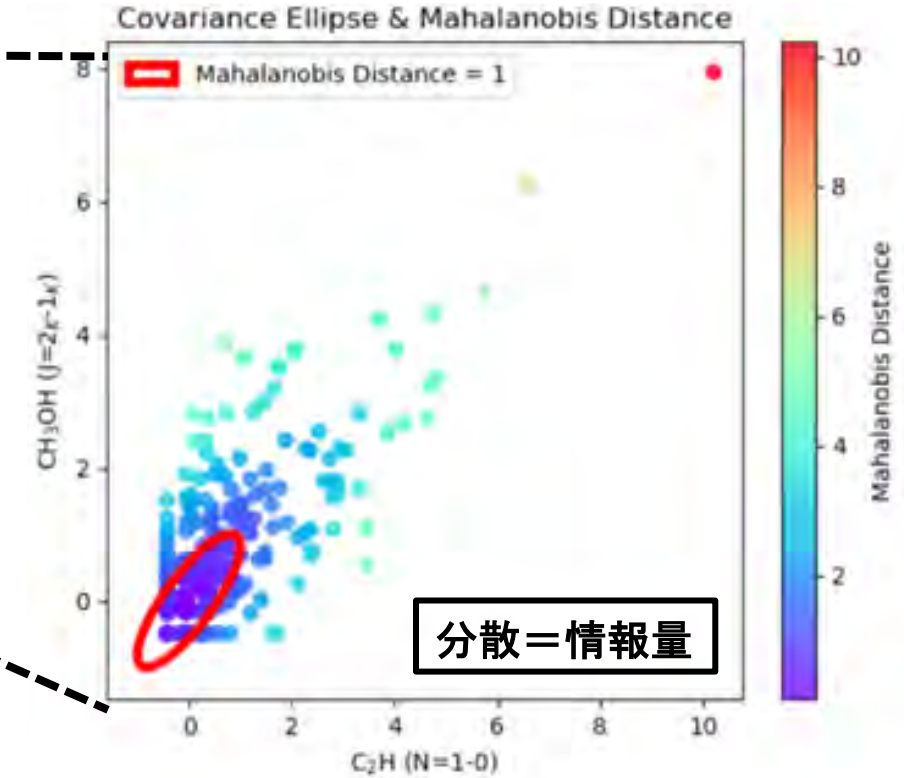
→輝線間の関係も定量的に調べるには？

共分散行列

・NGC 1068の13分子輝線に対する共分散行列



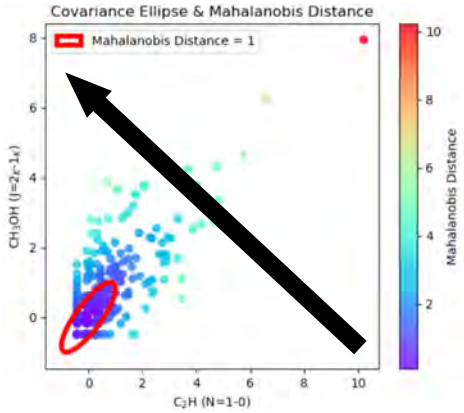




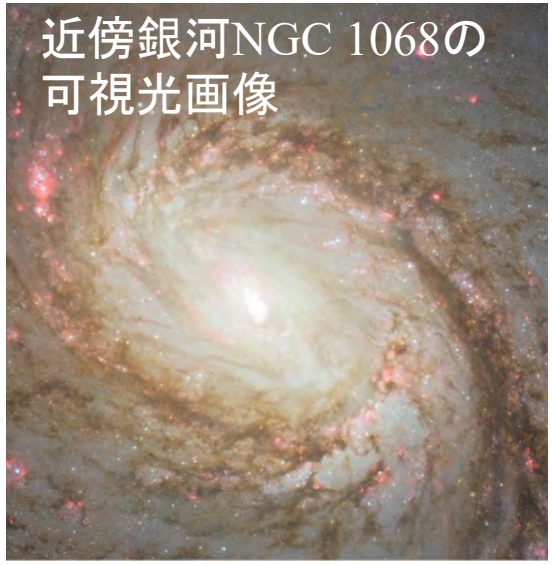
→相関関係を把握できる。
他の情報は？

主成分分析

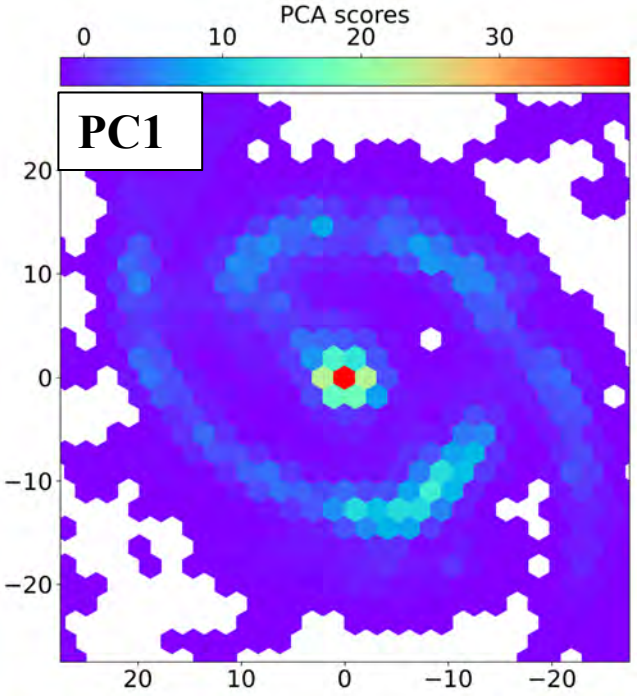
の固有値問題を解くと

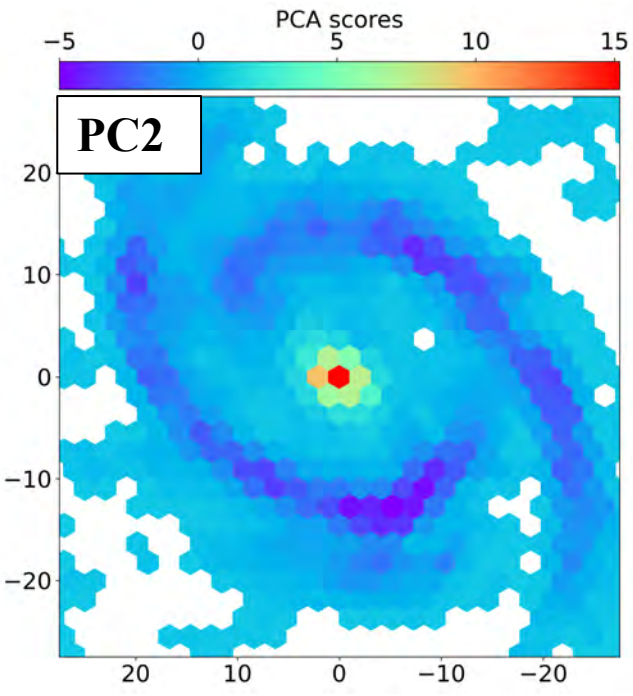


も把握できる。→ **主成分分析(PCA)**



13輝線に対する PCA

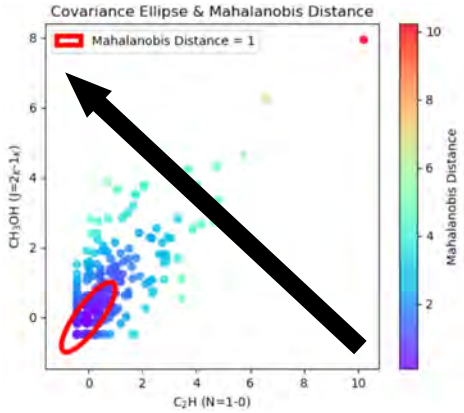




(Okubo et al. 2025)

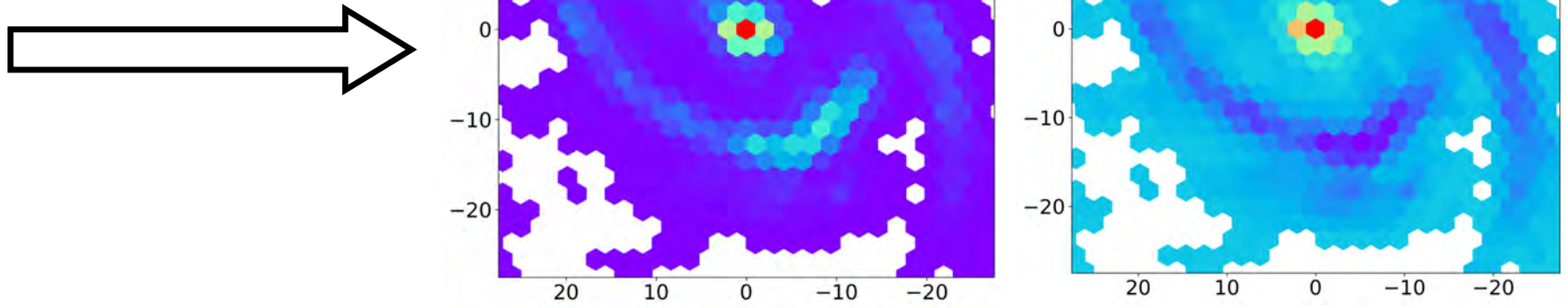
主成分分析

の固有値問題を解くと



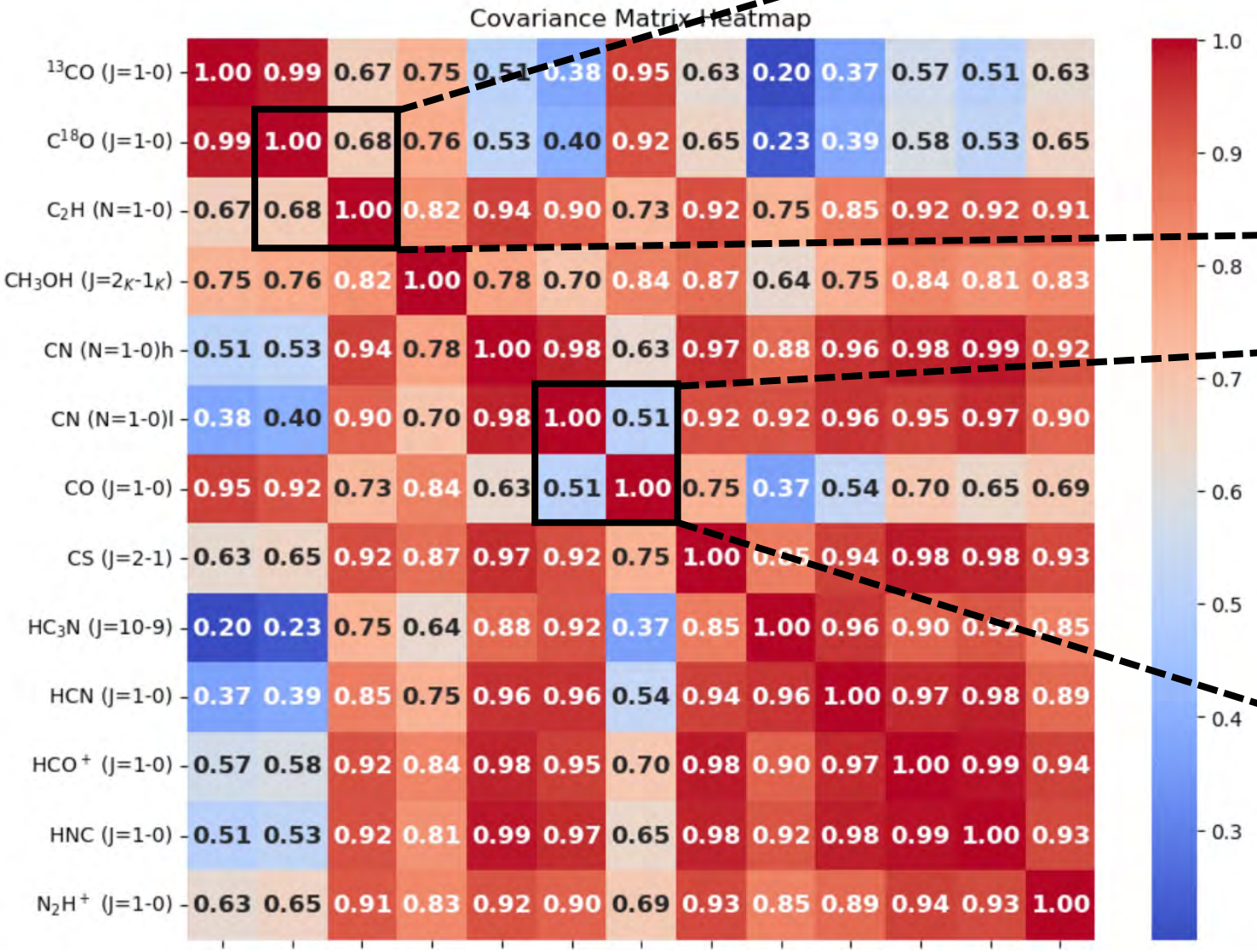
も把握できる。→ **主成分分析(PCA)**

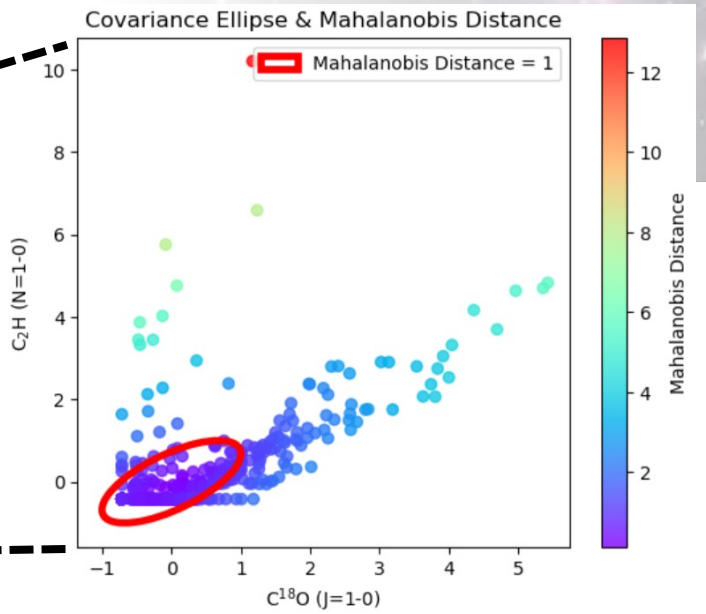
これで解決？

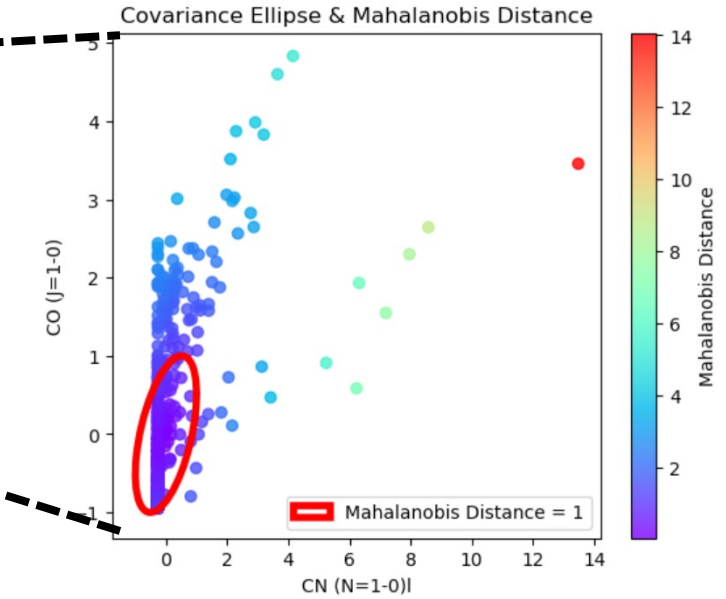


(Okubo et al. 2025)

PCAの欠点

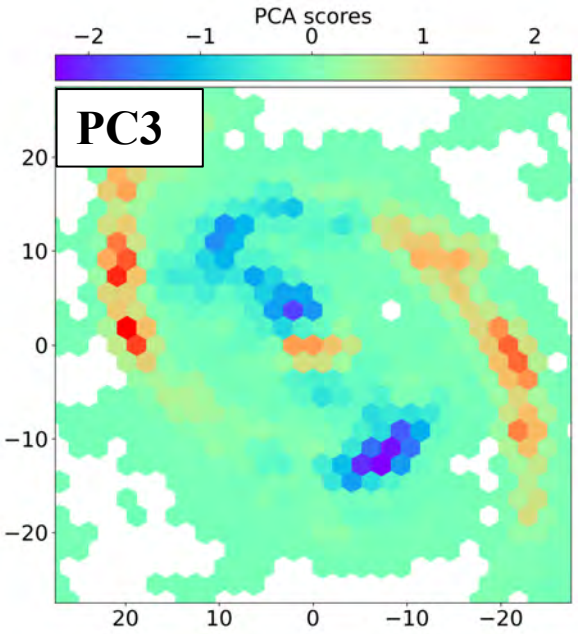


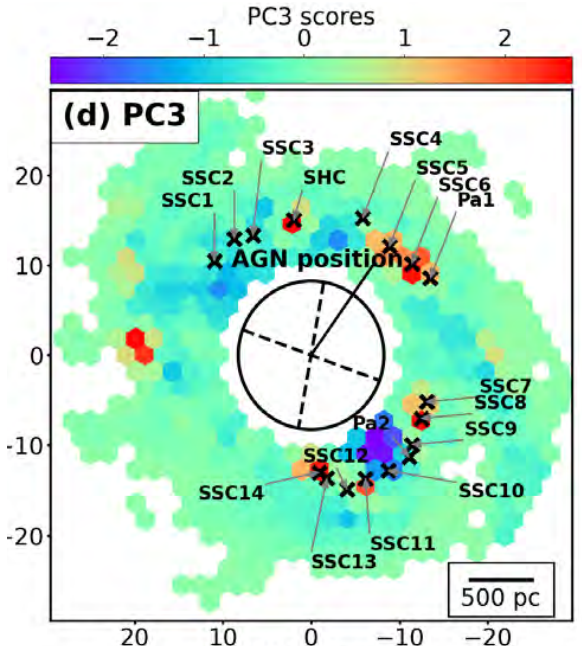




→ 非線形関係は把握できない。

PCAの欠点





(Okubo et al. 2025)

→ 複数の現象が混ざって抽出

→ 今まで見落としていた現象を発見？

『小さな情報を正確に抽出することで、天文学に新たな発見をもたらす』ことを示唆

“kernel PCA × kernel SHAP” による新たな解析手法を提示

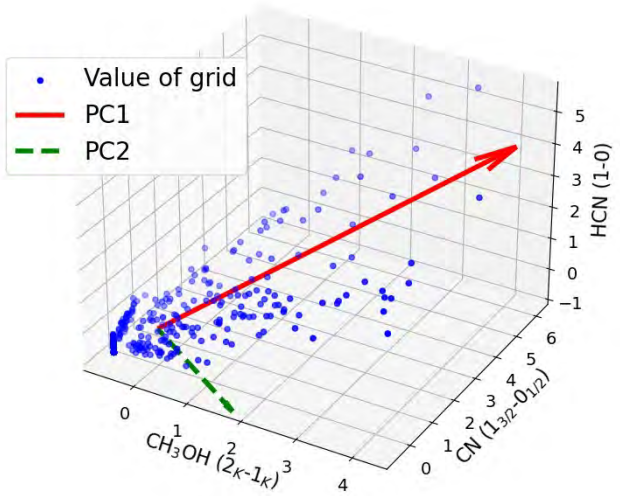
kernel PCAとは

kernel PCA: 非線形的な特徴抽出手法

データ行列が $X \in \mathbb{R}^{P\times N}$(P<N) の時(P:次元数、N:サンプル数)

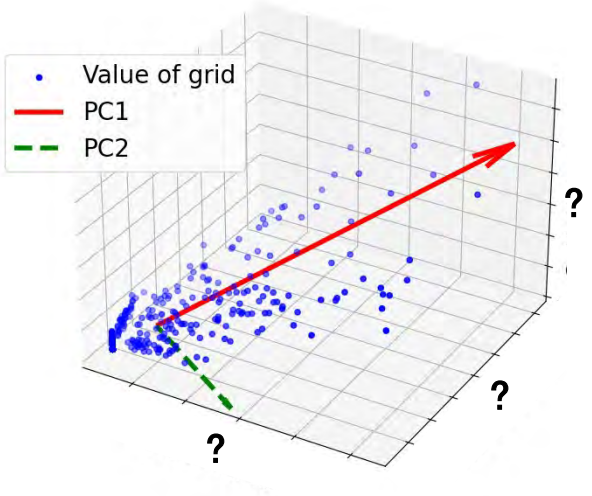
・PCAの場合

$x_i \in \mathbb{R}^P$



・kernel PCAの場合

$x_i \longmapsto \phi(x_i) \in \mathbb{R}^N$



kernel PCAの場合、PCA結果に対する輝線の寄与が分からない。

→ kernel SHAPで解決

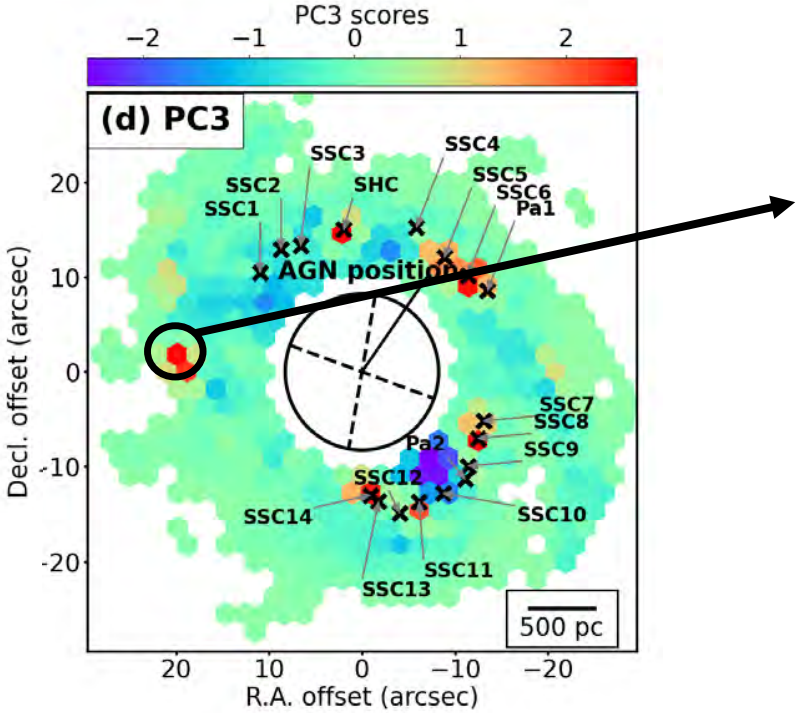
kernel SHAPとは

kernel SHAP:

・説明可能なAI (XAI)の1つ。

・個々のサンプルに対してインプットとアウトプットの関係を把握できる。

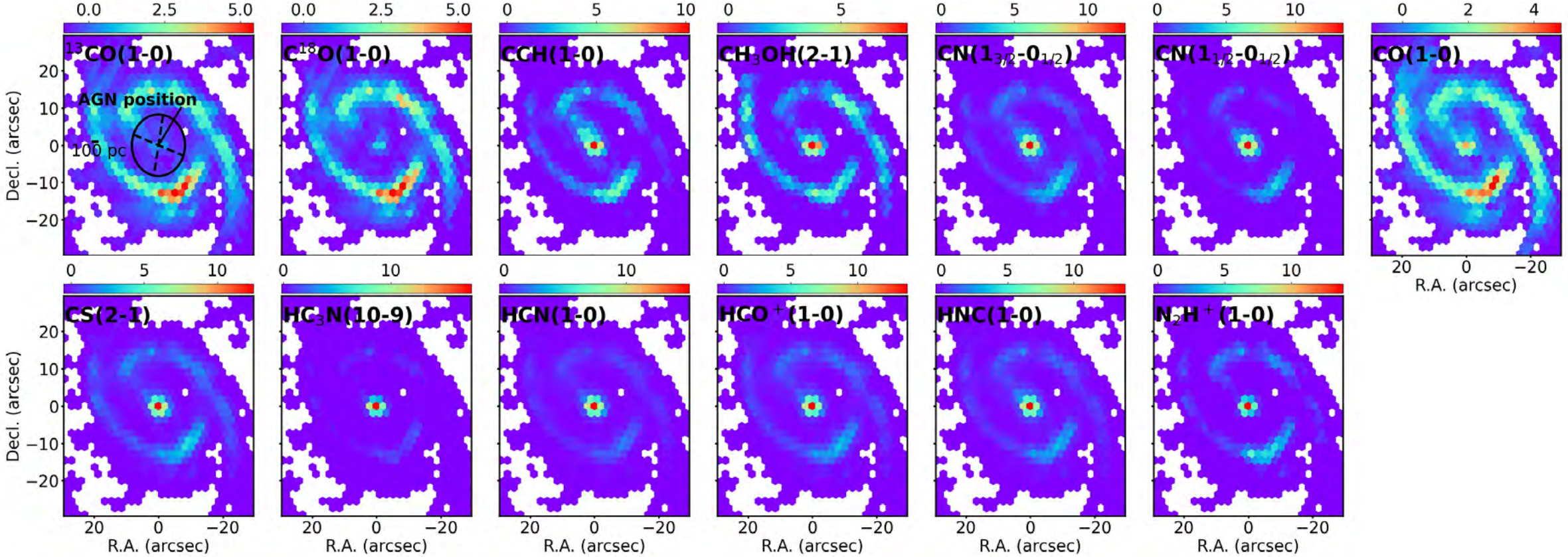
・SHAPのイメージ



score = line1 + line2 + line3 + … + lineP

データ

PCA結果と比較する必要あり。→ NGC 1068

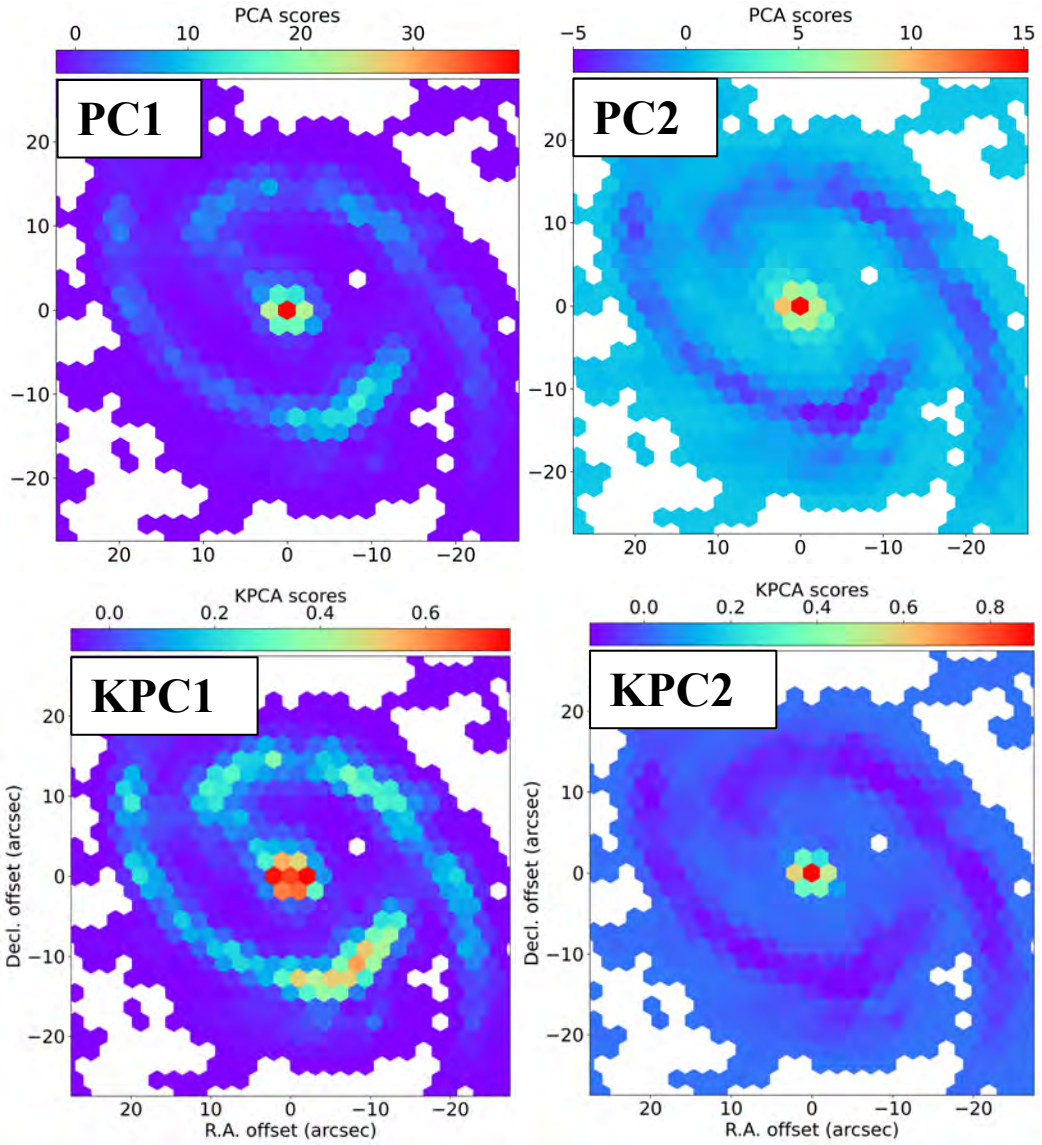


- ALMA Band3
- 空間分解能 : 150 pc
- S/N>4グリッドが60個以上の積分強度マップのみ使用→13分子輝線
- 標準化 (平均値0、分散1にする操作)

PCAとkernel PCA

PCA

kernel PCA

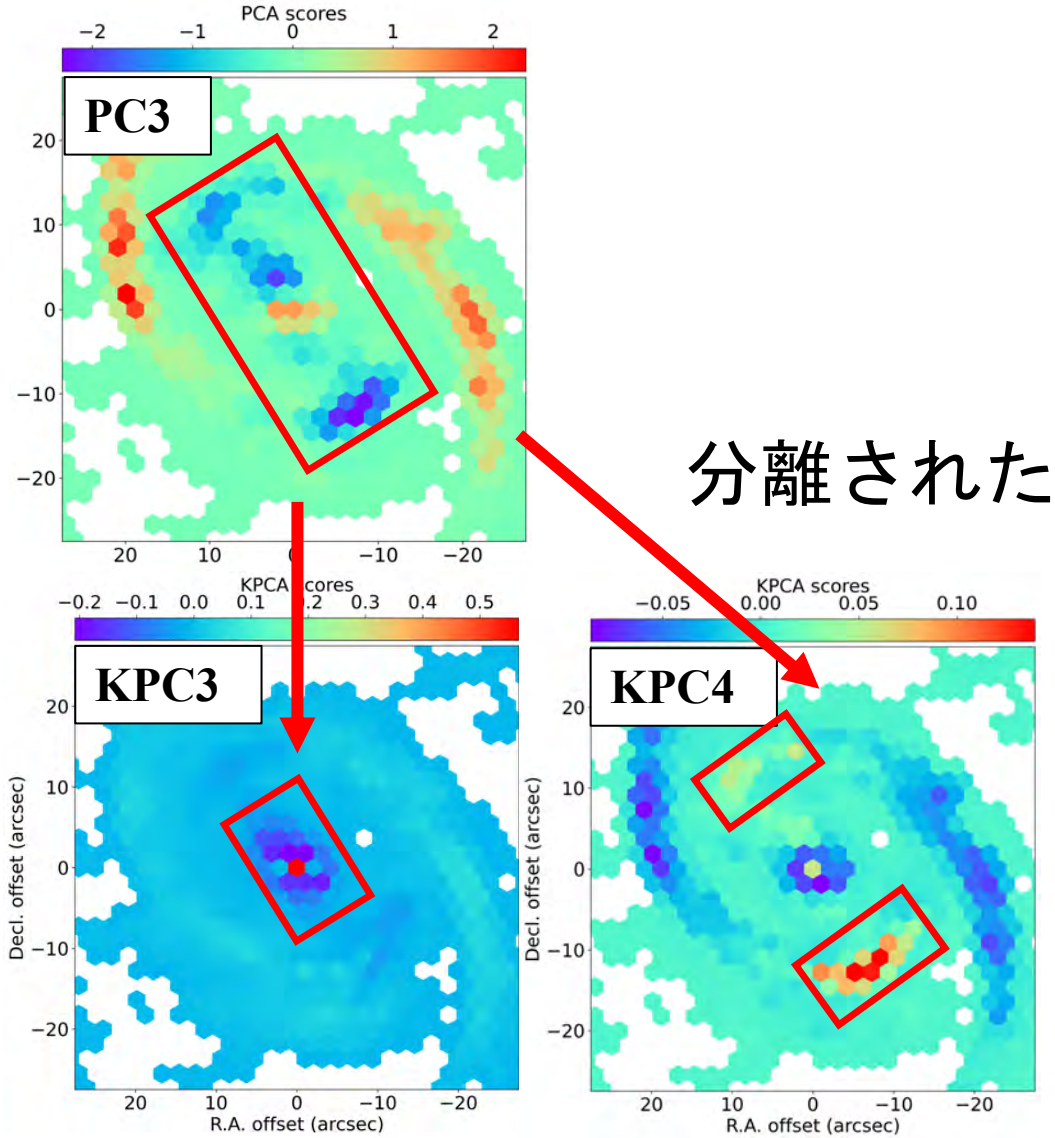


- kernel関数：RBFカーネル
- kernel係数：0.001

PCAとkernel PCA

PCA

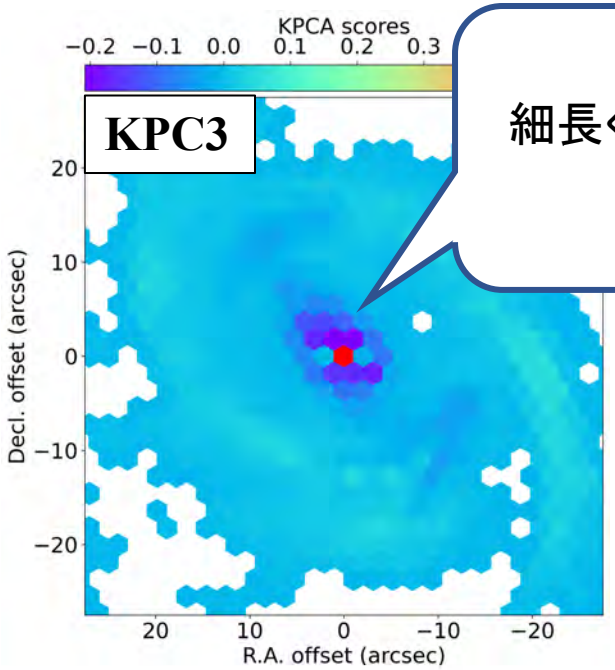
kernel PCA



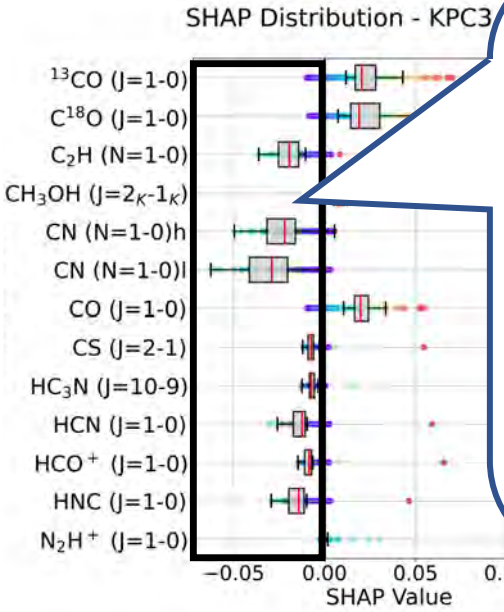
分離されたような見た目

KPC3の意味は？

本手法を使用する目的：分子輝線間の非線形性を可視化することで、未知の現象を抽出。
→ここでは、最低限のドメイン知識（分子輝線の性質）のみからKPC3 negativeの意味を推測。



細長く伸びているような
見た目



・特にCN, C_2Hの寄与が大きい。
・次いでHCN, HNC, HCO^+の寄与が大きい。
→ UV, 高密度ガスが関係してそう？

星形成？アウトフロー？
でもKPC3は150pcの分解能で、negative scoresは広がっている。
→星形成は考えにくい？アウトフローの可能性が高そう？

→『KPC3 negativeはアウトフローの特徴を抽出しており、それに伴う高密度化や化学組成の変化を表している』と予測

KPC3の意味は？

本手法を使用する目的：分子輝線間の非線形性を可視化することで、未知の現象を抽出。
→ここでは、最低限のドメイン知識（分子輝線の性質）のみからKPC3 negativeの意味を推測。

この結果を信用していいのか？

- kernel PCA：固有値問題を解くことによる非線形的な情報量の抽出
- kernel SHAP：ゲーム理論から派生した手法

→物理法則に基づいていない
→物理的な意味を持っているのか分からない

Decl. offset (arcsec)

→『KPC3 negativeはアウトフローの特徴を抽出しており、それに伴う高密度化や化学組成の変化を表している』と予測

考察 – 先行研究との比較

KPC3 negative

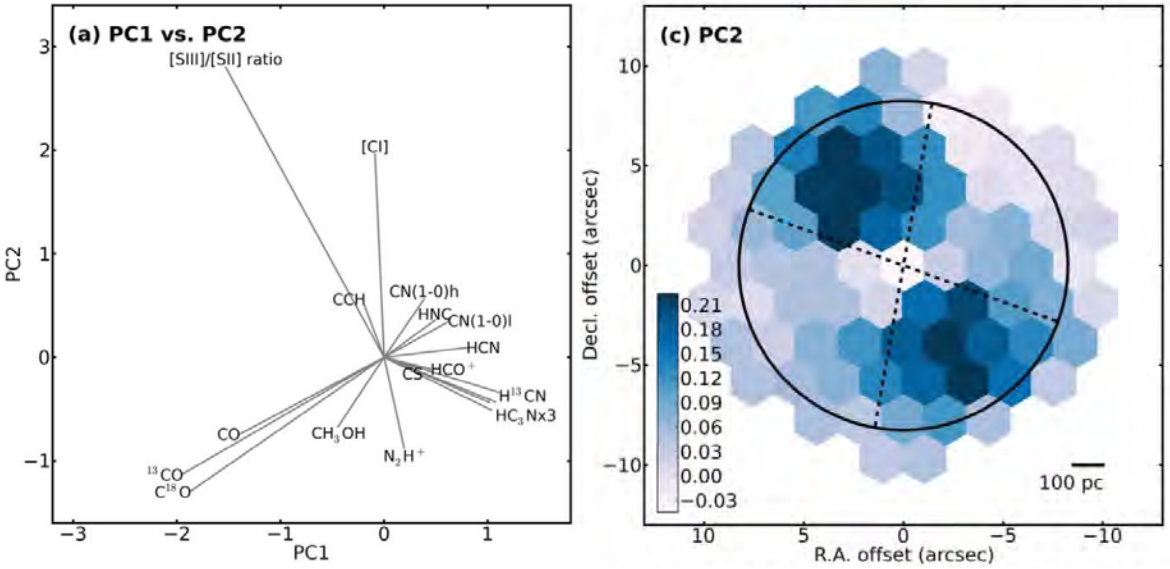
・アウトフローの存在を示唆: García-Burillo et al. (2014, 2017), Saito et al. (2022a, b), etc

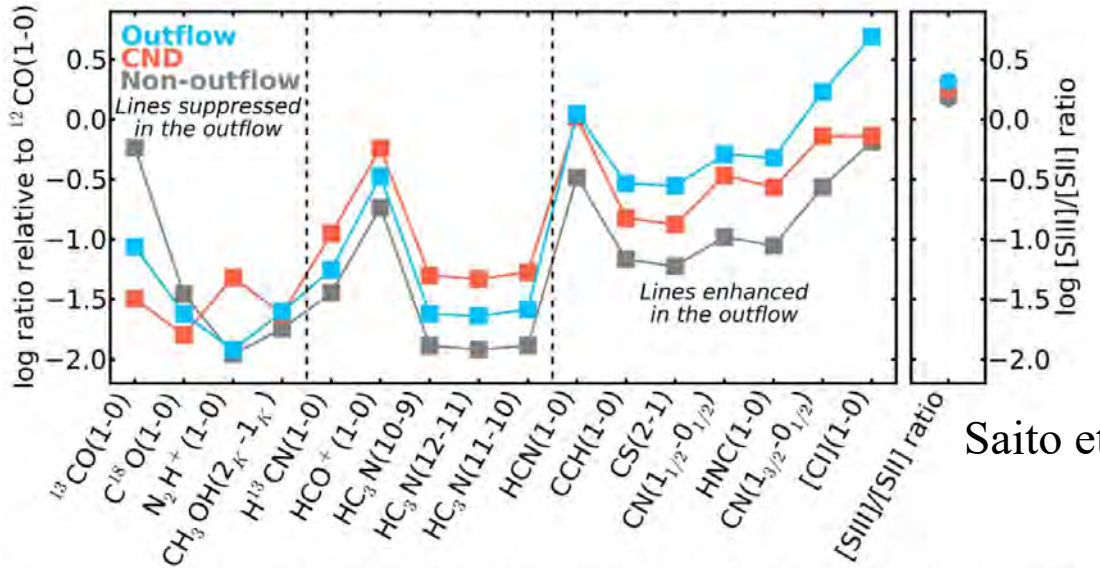
kernel SHAP

・アウトフローによる周囲のガスの高密度化を示唆: García-Burillo et al. (2014)

・アウトフローのUVによるCN, C_2H, HNC存在量の増加を示唆:
García-Burillo et al. (2017), Saito et al. (2022b)

→ KPC3 negative, kernel SHAPによる予測を支持



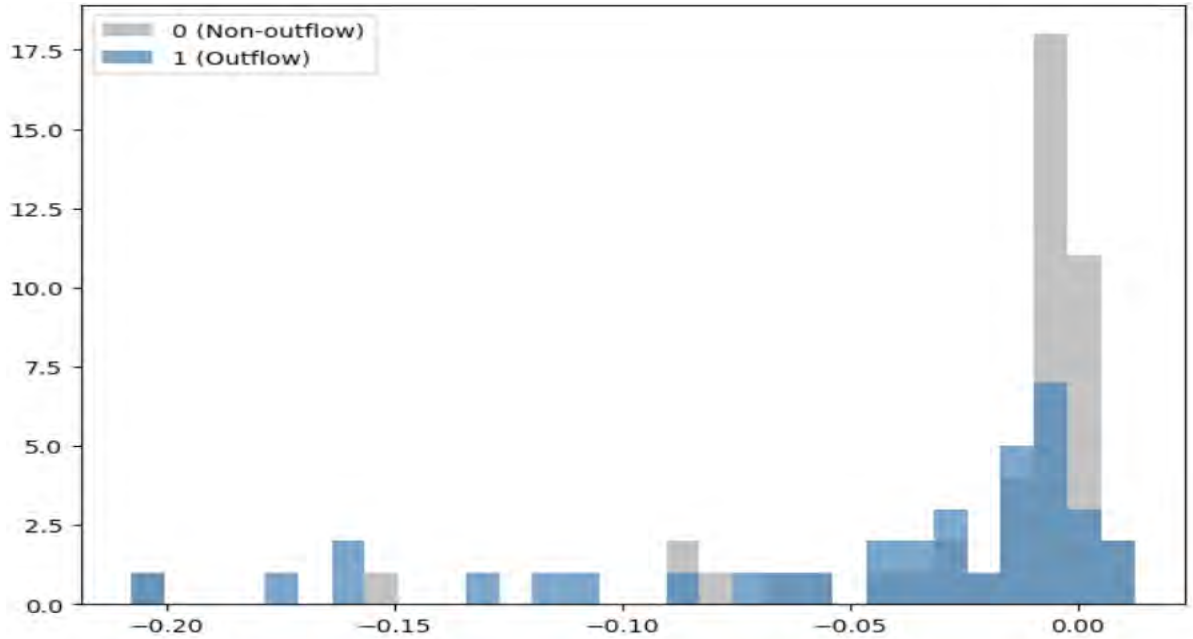


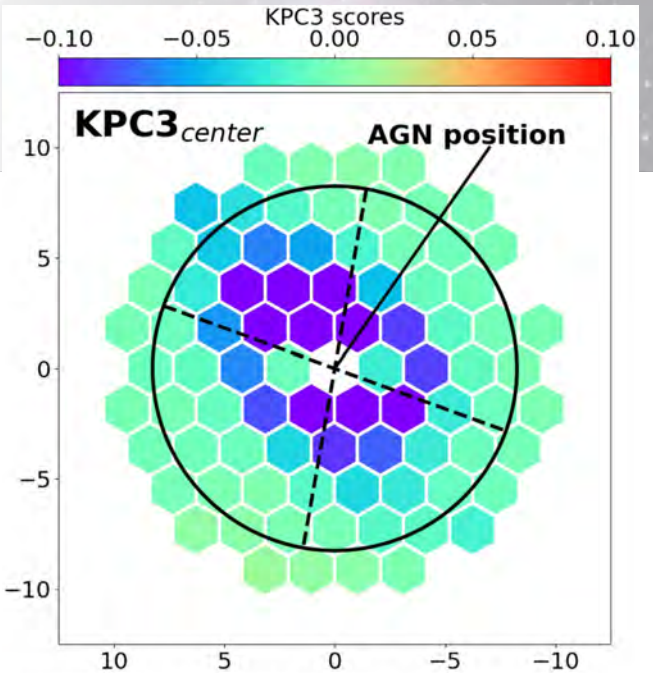
Saito et al. (2022b)

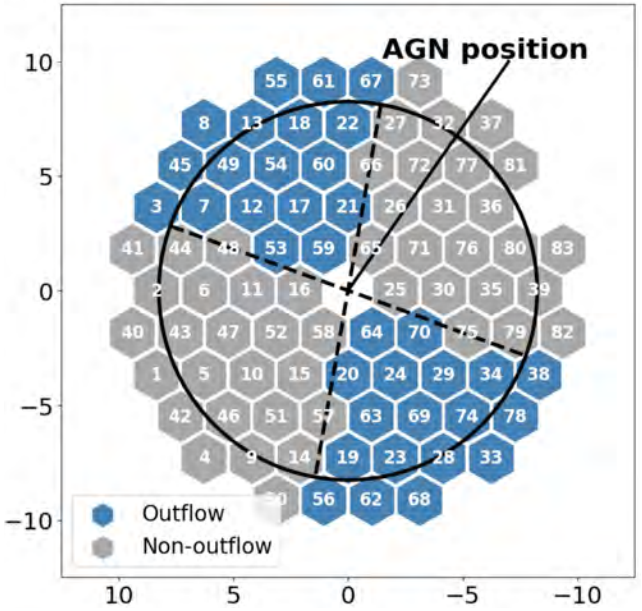
考察 - 2標本のt検定

KPC3は本当にアウトフローの特徴なのか？→ **2標本のt検定**

帰無仮説：KPC3_mean(Non-outflow) = KPC3_mean(Outflow)
対立仮説：KPC3_mean(Non-outflow) > KPC3_mean(Outflow)
有意水準：0.05







p値：0.01
→ 対立仮説を採用
→ KPC3 scoresはアウトフロー方向に分布
→ **KPC3 negative scoresはアウトフローの特徴を抽出**

考察 – 物理量との関係

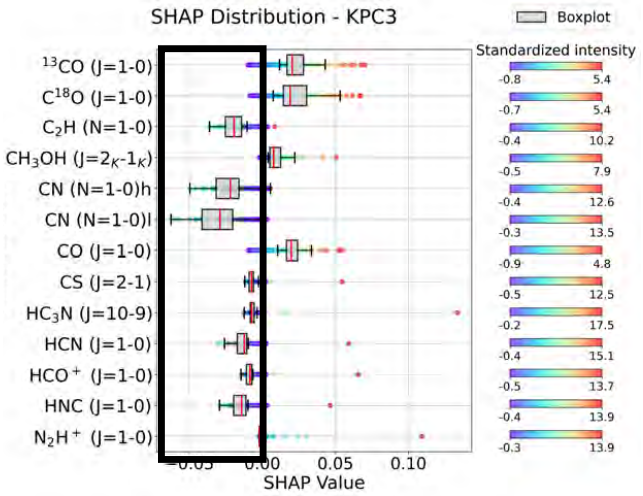
KPC3 negativeが物理情報を抽出していれば、物理量との相関を示すはず。
→ 最も基本的な物理量である**分子の柱密度**との相関をとる。

LTE解析結果

輝線名	^{13}CO	C_2H	CN	CS	HCN	HNC	HCO^+
相関	0.069±0.027	−0.628±0.023	−0.844±0.017	−0.782±0.006	−0.816±0.004	−0.842±0.003	−0.800±0.007

kernel SHAPが、KPC3 negativeに寄与すると判断した分子は、すべて強い相関を持つことが分かった。

→ kernel SHAPは物理的に意味を持つ情報を反映

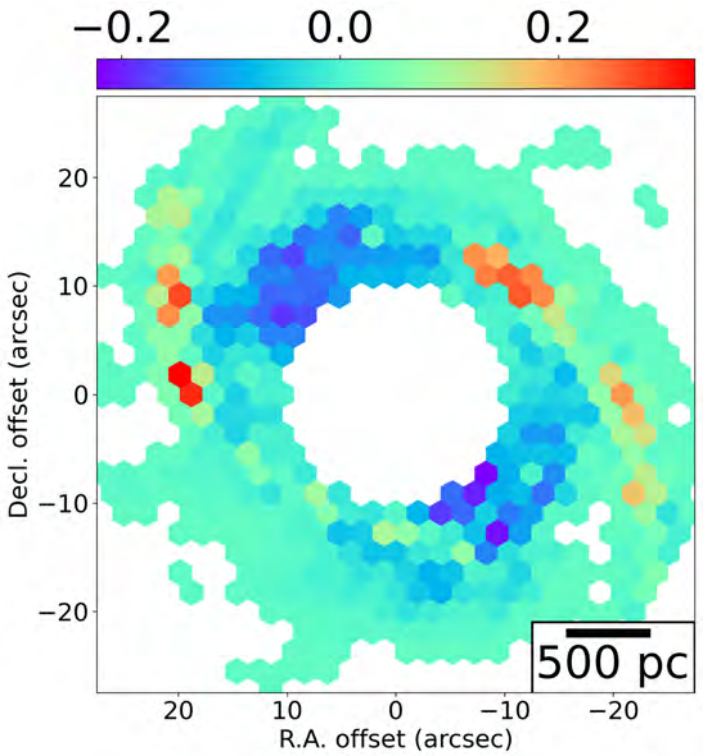


kernel PCAは、PCAでは抽出されなかった物理情報を抽出でき、kernel SHAPは、その結果を解釈できる。

まとめと今後

まとめ

- kernel PCA × kernel SHAPによる新たな手法を天文学に提示。
- 物理的に意味を持つ情報の抽出に成功。
- → 従来の解析では見逃されていた現象の抽出を示唆。

今後

- 本手法の有用性を示すため、見逃されていた現象の抽出を目指す。
- NGC 1068のアウトフローが星形成が盛んな領域に影響を及ぼしている可能性を初めて抽出。
- 今後は天文的観点から研究を実施。