

频率平均效应对SKA1-LOW成像影响的 仿真研究*

梁炜杰 谢扬帆 王锋[†] 邓辉 梅盈

(广州大学物理与材料科学学院 广州 510006)

摘要 平方公里阵列(Square Kilometer Array, SKA)射电望远镜的低频阵列SKA1-LOW将承担宇宙再电离(Epoch of Reionization, EoR)等关键科学研究任务. 频率平均计算可以有效降低数据量、提高成像灵敏度, 但所带来的“带宽效应”对SKA1-LOW成像的影响还缺少定量的分析和研究. 采用仿真方法, 对频率平均效应进行了深入研究. 通过对SKA1-LOW进行全阵列仿真、频率平均和洁化处理, 系统对比了平均多频道及单频道情况下的复原图, 分析两者相减后的残差图. 深入分析了频率平均对观测源强度的影响, 得到峰值强度在不同成像权重时随平均频道数变化的定量结果, 为SKA1-LOW实际观测中选择合适的频率平均频道数提供了参考依据, 对于SKA1-LOW开展科学研究具有重要的价值.

关键词 仪器: 干涉仪, 技术: 干涉, 方法: 解析, 方法: 数值
中图分类号: P164; 文献标识码: A

1 引言

平方公里阵列(Square Kilometre Array, SKA)^[1]是目前正在建设的具有重大意义的射电望远镜阵列项目, 其第1阶段(SK1)是射电天文学发展的重要内容. SK1由低频阵列(SK1-LOW)和中频阵列(SK1-Mid)构成, 旨在攻克天文学领域的重大科学难题. 其中, SK1-LOW部署于澳大利亚西部的沙漠地带^[2], 配备了131072个宽波段的双极化天线, 观测频率范围为50 MHz至350 MHz. 这些天线被划分为512个站点, 每个站点由256根天线组成.

SK1-LOW的建设是为了实现多个重要科学目标, 这些科学目标的达成将助力探索宇宙更多

的奥秘. 其中的高优先级目标(High-priority Science Objective, HPSO)包括对再电离时期(Epoch of Reionization, EoR)和宇宙黎明(Cosmic Dawn, CD)红移 $z < 27$ 时期的原子氢21 cm发射和吸收谱线的研究、脉冲星巡天以及脉冲星计时观测^[2]. 而实现这些科学目标离不开对天空进行多波段的观测, SK1-LOW不仅拥有着300 MHz的宽波段, 还具备极高的频率分辨率(单个频道带宽为5.4 kHz)^[3]. 如此窄的频道带宽虽提供了更精细的频谱数据, 但同样会产生海量数据, 给后续数据处理带来了极大的挑战.

随着SKA开始AA0.5阶段的调试, AA2即将动工. SKA的科学观测提案申请已经逐渐提上议题. 针对不同的科学目标, 如何选择合适的观测与数

2025-01-23收到原稿, 2025-03-21收到修改稿

*国家自然科学基金项目(12433012、12373097、12573094), 广东省自然科学基金项目(2024A1515011503), 广州大学研究生创新能力培养计划(JCCX2024-011), 中国科学院“西部之光”人才培养计划(xbzbz-zdsys-202410)资助

[†]fengwang@gzhu.edu.cn

据处理模式,是未来开展SKA科学研究的一个重要前提.为缓解数据处理的负担,通常需要对可见度(visibility)数据在频率轴上进行平均化处理,此操作称为频率平均.然而,频率平均在减少数据量、提升灵敏度的同时,会损失部分频谱结构,且会导致非相位中心源在径向上出现拖长和峰值强度下降的现象,此现象被称为“带宽拖影”或“带宽效应”(bandwidth smearing or decorrelation)^[4-5].而频率平均所带来的带宽效应是确定观测方案时需要考虑的一个关键问题.SKAO官网中给出一系列在线工具来帮助科学家确定SKA观测参数.其中灵敏度工具^①用来帮助科学家在给定不同的观测频率、不同的频率平均情况下计算SKA望远镜的灵敏度^[6].但这一工具无法给出非相位中心源的流量变化和半高全宽的变化情况,因此无法评估频率平均对成像的影响.事实上,就SKA的成像观测而言,只考虑灵敏度是不完整的.科学家需要重点考虑在多个频率平均时如何在灵敏度和带宽效应两个方面的折衷问题.

考虑到频率平均引起的带宽拖影在成像(脏图)时产生的影响以及退卷积可能无法消除此影响,对由频率平均所引起的图像变化进行定量分析,能为后续确定SKA的观测模式与数据处理模式提供关键指导.本文系统开展了SKA多频观测和数据处理中的带宽拖影问题研究工作,第2节概述了本次研究的具体步骤,第3节概述了本研究的结果及其分析,第4节进行了总结和展望.

2 基于仿真的带宽效应研究

本文将基于定量的仿真实验分析频率平均所引起的带宽拖影对成像的影响.在其他观测条件和退卷积参数保持一致的前提下,本文将对比并分析平均多频道及单频道(未进行频率平均)情况下的复原图.从仿真数据集到复原图的具体步骤在下文进行详细阐述,具体流程如图1所示,其中ms表示测量集(Measurement Set),WSCLEAN为退卷积软件,全称为W-Stacking Clean.

^①<https://sensitivity-calculator.skao.int/low>

^②<https://ska-telescope.gitlab.io/sim/oskar/>

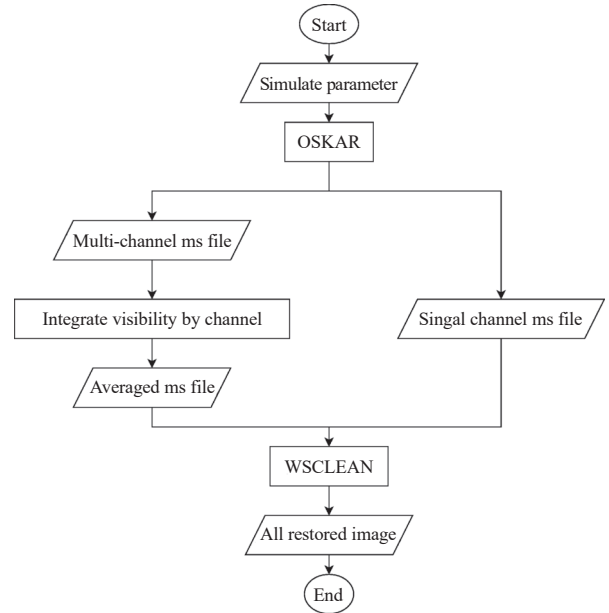


图 1 从仿真到复原图的流程图

Fig. 1 The flowchart from simulation to restored image

2.1 仿真

首先,我们使用OSKAR软件^②对SKA1-LOW的全阵列进行仿真.仿真参数为:观测时间为上中天的前2 h和后2 h、积分时间为1 s、间隔时间为1 h、观测频率为151 MHz,一个频道的带宽为5.4 kHz,相位中心的赤经(RA)、赤纬(DEC)为 $[358.14^\circ, -35.47^\circ]$.为有效比较多个频道频率平均后和单个频道不平均的区别,频率平均的频道倍数为1、4、8、12、16、20、24、28、32.

对于仿真天体的选择,我们选取了由默奇森广角射电阵(Murchison Widefield Array, MWA)^[8]公开发布的默奇森广角阵河内外全天巡天(Galactic and Extragalactic All-sky MWA survey, GLEAM)^[9]中的部分数据,共81个点源,其频谱指数在 $[0.22, -1.57]$ 区间.在151 MHz的频率下,81个源皆在1 Jy以下,最大流量为0.856 Jy且平均流量为0.23 Jy.以32个频道的可见度数据为例,平均前的 uv 覆盖(uv coverage)如图2所示. uv 指

代 uvw 坐标系中的 u - v 面, 其中 uvw 坐标系为空间频率域的坐标系.

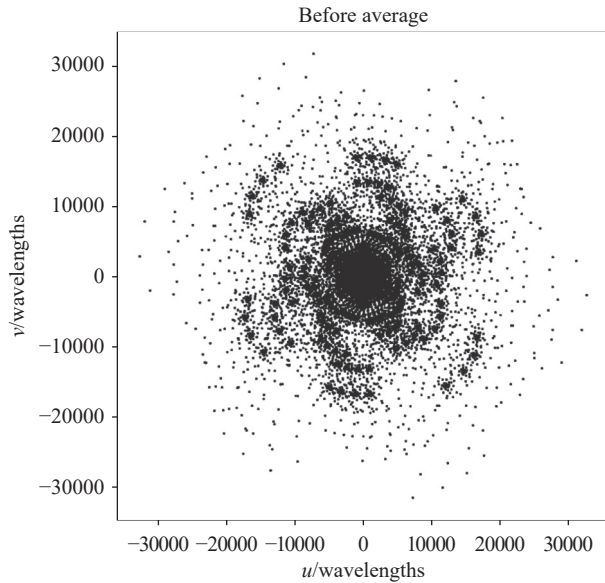


图2 SKA1-LOW全阵列在32个频道的 uv 覆盖

Fig. 2 32 channel uv coverage of the SKA1-LOW array

2.2 频率平均

基于仿真的可见度数据(以测量集数据格式存储), 对其进行频率平均计算. 我们基于团队开发的RASCIL2 (The Radio Astronomy Simulation, Calibration, and Imaging Library Ver 2.0)软件包, 进一步开发了实验分析代码. 具体步骤如下:

(1)调用 `create_visibility_from_ms` 函数将ms文件读入, 并自定义一个继承`xarray.Dataset` 类的子类`Visibility`用于封装传入信息;

(2)调用`integrate_visibility_by_channel` 函数对`Visibility`在频率轴上进行平均化处理, 并同步更新频率信息, 从而根据已知的以米为单位的 uv 坐标来更新以波长数为单位的 uv 坐标;

(3)调用 `export_visibility_to_ms` 函数将平均处理后的`Visibility`导出为ms文件.

2.3 成像与洁化

在频率平均计算后的可见度数据的基础上,

我们使用了目前主流的射电干涉阵数据处理软件WSCLEAN^③[10-11] (W-Stacking Clean)进行不同权重的干涉成像以及退卷积计算. 由于仿真的是点源, 此处用的退卷积算法是Cotton-Schwab CLEAN算法[12].

调用WSCLEAN的具体参数如下: 图像大小为 8192×8192 个像素、一个像素大小为 $2''$ 、每次在主循环中的增益(gain)为0.6、迭代次数设置为50000次、阈值为5倍标准差(σ). 为更加深入分析频率平均对复原图的影响, 使用了自然权重(natural weighting)和均匀权重(uniform weighting). 除此之外其他则用默认参数.

3 频率平均对成像质量的影响分析

为分析频率平均对复原图的影响, 本研究首先分析多个频道频率平均后的复原图与单个频道的复原图相减后的强度残差图. 为进一步分析频率平均对观测的影响, 本文将深入研究同一源在不同频道数平均后的强度变化情况.

3.1 图像复原结果对比分析

为分析频率平均导致的成像区别, 我们评估多个频道频率平均后的复原图减去单个频道复原图的残差图. 由于篇幅限制, 此处仅展示部分对比结果. 在自然和均匀权重下残差图的一个点源的差异效果如图3所示. 此点源亮度为 0.13 Jy 、频谱指数为 -0.69 、且位于 $[356.08^\circ, -34.36^\circ]$.

从图3中可以看出, 多频道的频率平均相对于单个频道无论是自然权重还是均匀权重成像, 在退卷积计算后皆出现较大的差异. 通过对比均匀权重与自然权重的残差图, 可以明显发现均匀权重的残差范围较小. 但在自然权重下, 频率平均引起的差异范围较大. 而且随着平均频道数的增加, 源的差异范围逐渐变大. 这表明在更多频道数平均的情况下, 采用自然权重成像时, 频率平均对源的影响范围更大.

与此同时, 可以看到源的中心区域基本为负数. 这表明相对单个频道, 多个频道频率平均后源

③<https://gitlab.com/aroffringa/wscclean>

中心区域的强度是降低的. 为进一步分析频率平均对观测源在成像中的影响, 本文后续将深入分

析其对源的峰值强度的影响.

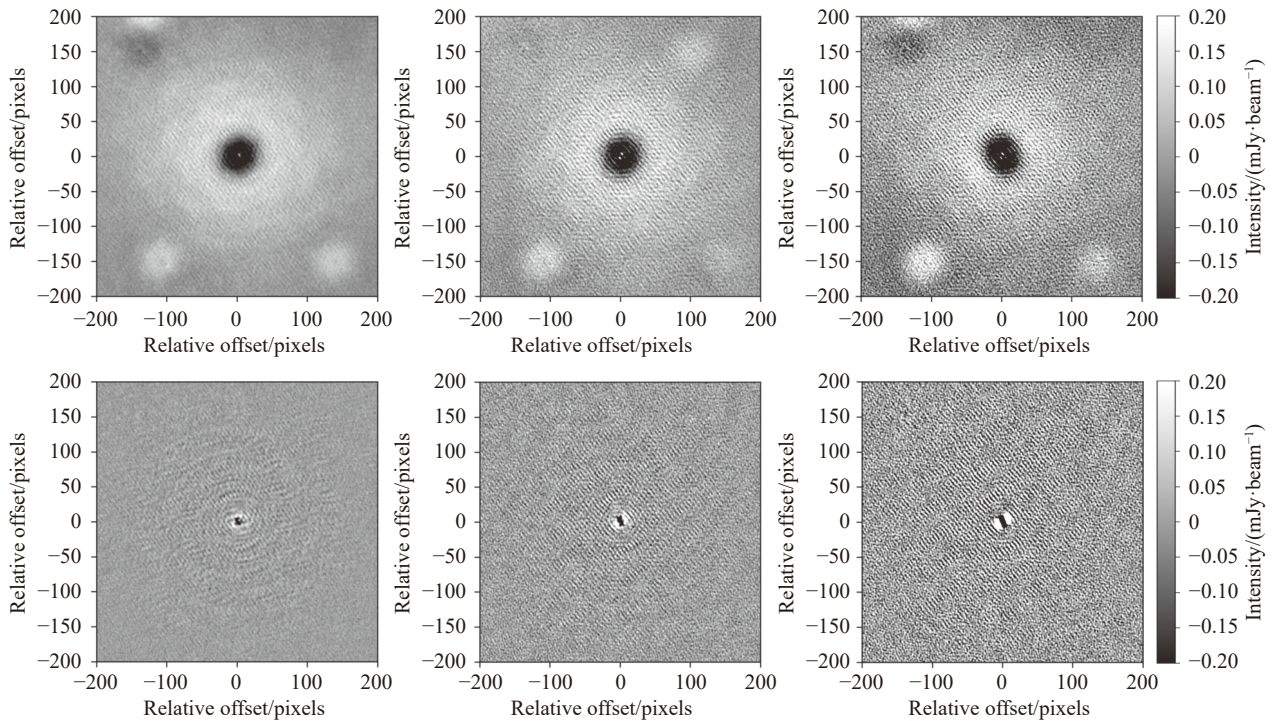


图 3 一个点源在自然(上行)和均匀(下行)权重下, 8 (左)、16 (中)、32 (右)个频道与1个频道的残差图

Fig. 3 Residual graph for 8 (left), 16 (middle), and 32 (right) channels versus 1 channel for a point source with natural (top row) and uniform (bottom row) weighting

3.2 频率平均对观测源强度的影响分析

为进一步分析频率平均对观测源的影响, 选取了在自然权重和均匀权重下复原图中的3个不同位置的点源的峰值强度变化作为研究对象, 结果如图4所示. 为更好地展示峰值强度的变化情况, 此图使用了点源在多频平均的峰值强度(I)与单频的峰值强度(I_0)的比值作为纵轴. 为更好地说明源的位置, 此图使用了点源相对相位中心的距离(θ)作为标签.

从图4中可明显看出随着源与相位中心距离增加以及平均频道数的增加, 无论是自然权重还是均匀权重, 源的峰值强度都在减弱. 并且自然权重相对于均匀权重, 源的峰值强度下降较慢.

基于此分析结果, 可在频率平均时选择频道

数提供参考作用. 例如: 对于SKA1-LOW全阵列的自然权重成像(完成退卷积后), 经过32个频道平均, 一个距离相位中心大约为 2° 的源, 可造成大约相对于一个频道成像1.32%峰值强度损失. 若是均匀权重成像, 则会造成5.64%峰值强度损失.

3.3 频率平均对图像中其他参数的影响分析

除了上述对源的峰值强度分析外, 我们还进一步考虑了有噪声影响的情况. 在有噪声时不同频率平均、不同成像权重的复原图中, 对观测源总流量、图像噪声水平以及源的波束大小进行分析研究. 我们使用了PyBDSF (Python Blob Detector and Source Finder)工具^④对复原图进行搜索统计, 结果如下. 这些分析结果对在频率平均时选择频道数也有参考作用.

^④<https://github.com/lofar-astron/PyBDSF>

图5展示了总流量的变化情况, 使用点源在多频平均的总流量(S)与单频的总流量(S_0)的比值作为纵轴. 可以看出随着频道的增加、不同成像权

重以及不同位置的点源, 总流量的值非常接近, 上下浮动仅约为0.5%. 这个差异可能是源经过退卷积后产生的偏差导致.

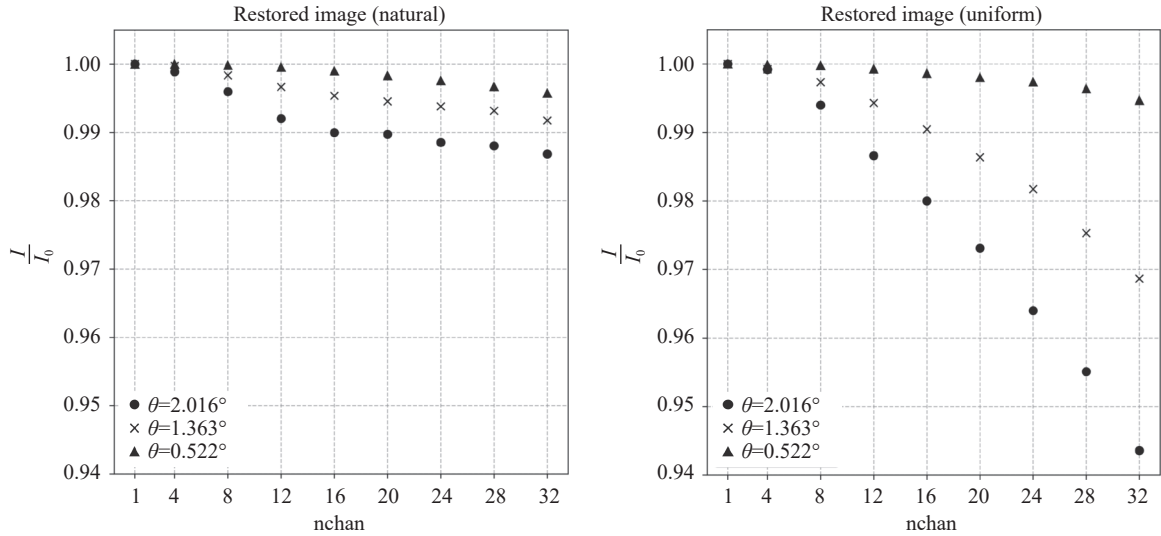


图 4 3个不同位置的点源在自然权重(左)和均匀权重(右)下复原图的峰值强度随平均频道数的变化, 其中nchan表示被平均的频道数.

Fig. 4 The peak intensity of point sources at three different locations in the restored image under natural weight (left) and uniform weight (right) varies with the average number of channels, where nchan denotes the average number of channels.

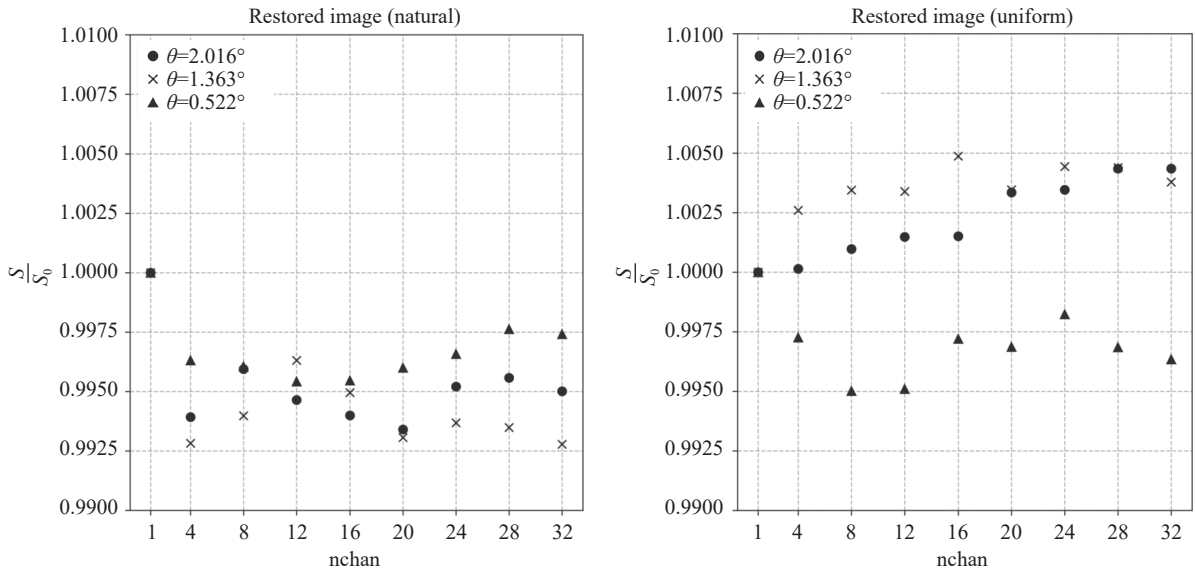


图 5 3个不同位置的点源在自然权重(左)和均匀权重(右)下复原图的总流量随平均频道数的变化

Fig. 5 The total flux of point sources at three different locations in the restored image under natural weight (left) and uniform weight (right) varies with the average number of channels

在图6中展示了不同成像权重下复原图的噪声水平. 从图中可以看出, 随着频率平均的频道数的增加, 无论是自然权重还是均匀权重的复原图, 噪声水平皆在下降, 且均匀权重的噪声水平皆大于自然权重的.

图7展示了波束长轴的变化情况, 此图使用了角秒作为点源波束的长轴单位, 并以此当作纵轴. 从图中可以看出在自然权重下, 3个点源的波束长轴几乎没有变化. 而在均匀成像权重下, 3个点源波束的长轴皆有增长的趋势, 并且位于距离相位中心约 2° 的点源波束长轴增长程度最大. 对于自然权重中波束长轴的变化异常, 这可能是源经过退卷积后产生的偏差导致.

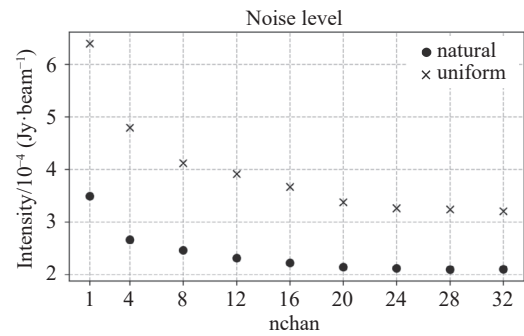


图 6 在自然权重和均匀权重下复原图的噪声水平随平均频道数的变化

Fig. 6 The noise level in the restored image under natural weight and uniform weight varies with the average number of channels

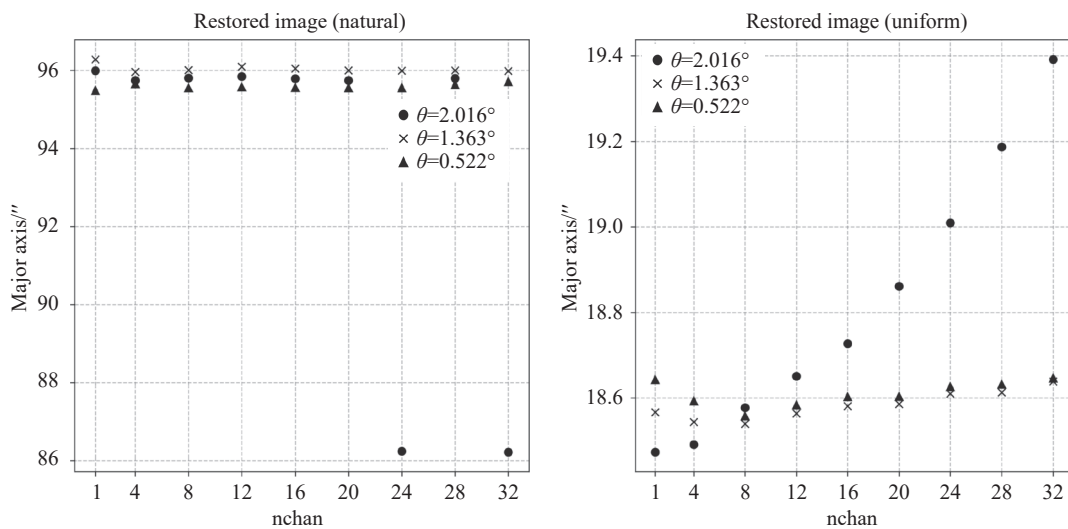


图 7 3个不同位置的点源在自然权重(左)和均匀权重(右)下复原图的长轴随平均频道数的变化

Fig. 7 The major axis of point sources at three different locations in the restored image under natural weight (left) and uniform weight (right) varies with the average number of channels

4 结论

本文基于OSKAR软件, 利用GLEAM星表部分源作为仿真数据集, 使用SKA1-LOW的全阵列进行了数据仿真. 并利用RASCIL2对仿真数据进行频率平均. 通过WSCLEAN软件对经过频率平均的仿真数据进行退卷积处理和最终成像. 通过对比多频道频率平均和单频道不平均的成像结果, 分析频率平均引起的带宽拖影对退卷积的影响.

并深入分析此影响对观测源强度的影响, 得到不同位置的点源在不同成像权重下峰值强度随频道数增加的变化程度. 除此之外, 还分析了图像中观测源的总流量、波束大小以及图像的噪声水平随平均频道数的变化情况.

研究成果表明, 频率平均对复原图的源结构和峰值强度都具有显著影响. 其中均匀权重相对于自然权重对源的影响范围较小, 但对源的峰值强度影响较大. 并且得到了点源的峰值强度、总

流量、波束大小以及图像的噪声水平随条件变化的定量结果, 基于这些结果, 在实际观测中, 可参考这些结果并结合具体需求选择合适的频率平均频道数, 以优化成像质量. 总体来看, 本研究为频率平均在射电天文观测中的应用提供了实践指导, 对未来开展SKA1-LOW的科学数据处理有较好的参考价值.

需要指出的是, 本工作仍存在一定的局限性. 例如, 本研究未考虑天线的主波束效应等实际因素, 与真实观测环境存在着一定的差距. 未来的研究可以进一步引入这些实际因素, 以提高仿真实验的准确性.

参考文献

- [1] Dewdney P E, Hall P J, Schilizzi R T, et al. *IEEEP*, 2009, 97: 1482
- [2] Labate M G, Braun R, Dewdney P, et al. 2017 XXXIInd General Assembly and Scientific Symposium of the International Union of Radio Science (URSI GASS). Montreal, QC. *IEEE*, 2017: 1-4
- [3] Dewdney P E, Labate M G, Swart G, et al. SKA1 Design-Baseline Description. SKA-Tel-SKO-0001075-rev2, 2022
- [4] Bridle A H, Schwab F R. *Synthesis Imaging in Radio Astronomy II*. San Francisco, Calif, United States of America: ASP Conference Series, 1999, 180: 371
- [5] Moran J M, Swenson G W. *Interferometry and Synthesis in Radio Astronomy (3rd)*. Cham, Switzerland: Springer Nature, 2017: 309-387
- [6] Sokolowski M, Tingay S J, Davidson D B, et al. *PASA*, 2022, 39: e015
- [7] Mort B J, Dulwich F, Salvini S, et al. 2010 *IEEE International Symposium on Phased Array Systems and Technology*. Waltham, MA, USA: *IEEE*, 2010: 690-694
- [8] Tingay S J, Goeke R, Bowman J D, et al. *PASA*, 2013, 30: e007
- [9] Hurley-Walker N, Callingham J R, Hancock P J, et al. *MNRAS*, 2017, 464: 1146
- [10] Offringa A R, McKinley B, Hurley-Walker N, et al. *MNRAS*, 2014, 444: 606
- [11] Offringa A R, Smirnov O. *MNRAS*, 2017, 471: 301
- [12] Schwab F R, Cotton W D. *AJ*, 1983, 88: 688

Simulation Analysis of Frequency Averaging Effects on SKA1-LOW Image

LIANG Wei-jie XIE Yang-fan WANG Feng DENG Hui MEI Ying

(School of Physics and Materials Science, Guangzhou University, Guangzhou 510006)

ABSTRACT The SKA1-LOW of the Square Kilometer Array (SKA) will perform key scientific tasks such as studying the Epoch of Reionization (EoR). Frequency averaging can reduce visibility data volume and enhance sensitivity, but it introduces “bandwidth smearing”. The quantitative impact of this effect on SKA1-LOW image remains unexplored. This study employs simulation to investigate frequency averaging. By simulating the SKA1-LOW full array, applying frequency averaging and cleaning processing, we systematically compare restored images from averaged multi-channel and single-channel data, analyze their residual after subtraction, and evaluate the effect on observed source intensity. Quantitative results are obtained for peak intensity variation with averaged channel count under different imaging weights. This research provides guidance for selecting optimal channel numbers in SKA1-LOW observations, offering significant value for SKA1-LOW science.

Key words instrumentation: interferometers, techniques: interferometric, methods: analytical, methods: numerical