

# DroPS 说明文档中文版

黄志琦

huangzhq25@mail.sysu.edu.cn

支撑项目：国家重点研发计划 2020YFC2201602

2025 年 11 月

# Contents

<b>1 绪论</b>	<b>3</b>
1.1 原初功率谱	3
1.2 CMB B 模式极化和张标比 $r$	3
1.3 地面小口径 CMB 望远镜	4
<b>2 软件使用说明</b>	<b>5</b>
2.1 安装 DroPS	5
2.1.1 安装系统环境	5
2.1.2 建立 Python 虚拟环境	6
2.1.3 安装 python 依赖	7
2.1.4 修改 pysm3 包	7
2.2 参照数据组	8
2.2.1 生成一个时序数据过滤的模拟器	8
2.2.2 生成参照数据组	9
2.3 分析天图	11
2.3.1 单张模拟观测天图的分析	11
2.3.2 多张模拟观测天图的分析	12
2.3.3 真实观测数据上的应用	13
2.3.4 和文献中同类软件比较	14
2.4 天图成分分离	14
2.4.1 基于似然函数的天图成分分离方法	14
2.4.2 可视化	18
<b>3 软件技术细节</b>	<b>19</b>
3.1 天图制作和时序数据过滤	20
3.2 噪声模型	21
3.3 CMB 模型	22
3.3.1 宇宙学	22
3.3.2 去透镜化	22
3.3.3 宇宙学参数限制	23
3.4 前景模型	24
3.4.1 尘埃热辐射和同步辐射的基本模型	24
3.4.2 前景能谱的空间变化: 矩展开方法	25
3.4.3 前景频率去相关: 泰勒展开方法	28
3.5 带功率的相关矩阵	29
3.6 天图成分分离的似然函数	32

# 1 绪论

## 1.1 原初功率谱

根据目前标准的宇宙学模型，宇宙微波背景辐射（CMB）和宇宙大尺度结构都源于原初宇宙的微小的，近似服从高斯统计的度规扰动。在线性微扰的层面上，这些原初度规扰动可以被分解为标量、矢量和（二阶）张量。因为矢量扰动会很快衰减，所以我们一般忽略它。宇宙学分析通常引入无量纲、各向同性的标量功率谱  $\mathcal{P}_S(k)$  和张量功率谱  $\mathcal{P}_T(k)$  来刻画原初的标量和张量扰动。这里的  $k$  是共动波矢的大小。

基于慢滚暴胀模型的预言，这两个功率谱通常参数化为幂率形式：

$$\mathcal{P}_S(k) = A_s \left( \frac{k}{k_{\text{pivot}}} \right)^{n_s-1}, \quad (1)$$

和

$$\mathcal{P}_T(k) = r A_s \left( \frac{k}{k_{\text{pivot}}} \right)^{n_t}. \quad (2)$$

现有的 CMB 实验已经对标量功率谱参数  $A_s$  和  $n_s$  给出了很强的观测限制 [1]. 在大多数暴胀模型中，张量功率谱的斜率  $n_t$  都很小，所以在实际数据分析中它或者被近似成零，或者按照单场慢滚模型的预言处理 ( $n_t = -r/8$ )。而剩下的刻画在选定的特征尺度  $k_{\text{pivot}}$  上的张量-标量-比（简称张标比）的参数  $r$ ，目前仍然没有观测到它的下限。现在最好的观测上限是  $r < 0.032$ （95% 置信度）[2]。虽然这个上限已经排除掉了一些暴胀模型，但是仍然有大量的模型和  $0 < r < 0.032$  兼容 [3].

## 1.2 CMB B 模式极化和张标比 $r$

在 CMB 中探寻原初引力波信号（测量  $r$ ）是现代宇宙学的最重要科学目标之一。这项探索聚焦于一个独特的可观测量：B 模式偏振图样。

CMB 光子在其散射早期宇宙中的自由电子时发生了偏振。这个过程在光线上印刻了方向性的偏好，从而产生了两种独特的模式：E 模式（类似于静电荷产生的电场<sup>1</sup>，标量扰动和张量扰动都可以产生）和 B 模式（类似于磁场<sup>2</sup>）。虽然原初等离子体中的密度涨落（标量扰动）能够产生 E 模式，并能通过引力透镜效应产生少量的 B 模式（即所谓的透镜化 B 模式），但它们无法产生原初 B 模式所特有的、大尺度的旋性图样。这正是原初引力波发挥作用之处。理论上产生于暴胀时期的原初引力波在早期宇宙中传播时，会周期性地拉伸和挤压空间，给等离子体施加了一种独特的、旋性的畸变，从而在 CMB 中产生了一种本质上具有旋转特性的偏振模式。

因此，CMB 在大角尺度上的 B 模式功率谱可以作为这些原初引力波的直接示踪物；探测到 CMB 中的大角尺度 B 模式功率谱几乎就等价于测到了原初引力波。因为

<sup>1</sup>这仅仅是一个类比。实际上  $E$  本身是一个标量场；和电场进行类比的极化矢量由  $E$  的二阶空间导数构造出来

<sup>2</sup>这也仅是类比。实际上  $B$  是一个赝标量场；和磁场类比的极化矢量由  $B$  的二阶空间导数构造出来。

$A_s \approx 2.1 \times 10^{-9}$  几乎是已经被确定的数, (2) 给出的原初引力波的功率谱大小直接正比于参数  $r$ 。用 CMB 的大角尺度 B 模式的观测数据限制  $r$ , 就成为窥探极早期物理, 探索引力的量子特性的重要窗口。

### 1.3 地面小口径 CMB 望远镜

地面小口径的望远镜 (SAT) CMB 实验, 包括 BICEP/Keck [4], AliCPT [5, 6, 7], Simons Observatory SATs [8, 9], 以及曾计划的 CMB-S4 SATs [10, 11]<sup>3</sup>等, 都有一个共同的目标: 从 CMB 中寻找原初引力波的独有信号。这些望远镜在 (大致在 30–300 GHz 范围内的) 多个频段观测前景相对较弱的部分天区。但即使是这些比较“干净”的天区, 前景依然远强于待测的信号。观测的原始数据将是 CMB 信号, 银河系前景, 大气辐射, 仪器噪声等的叠加。要把微弱的原初引力波信号从数据中提取出来 (即测量  $r$ ), 需要十分复杂的成分分离技术。这个文档对应的软件, DroPS (Deriving  $r$  from Power Spectra 的缩写), 就是为实现这个目标而设计。

---

<sup>3</sup>目前 CMB-S4 计划已取消。

## 2 软件使用说明

如果不涉及具体的从时序数据制作天图的流程，而直接从已有的天图出发限制  $r$ ，由于分析的是大角尺度（低分辨率）的 CMB 统计，所需要的计算资源并不多<sup>4</sup>。只有在处理真实数据，需要把 DroPS 连接到具体实验对应的天图制作数据管线时，才会涉及大量的计算资源。

因此，为了实现下面的教程的测试例子，用户只需准备一个有几百 GB 硬盘空间的个人电脑。

### 2.1 安装 DroPS

这个教程是在 Ubuntu 24.04 LTS 系统上测试的，但应该很容易通过微调在其他 linux 系统上实现。如果用户使用的是 Windows 或者 MacOS 系统，很多例子就可能需要较大的改动才能实现——如果环境的调试过于困难，一般的建议是安装一个 Ubuntu 24.04LTS 的子系统。

下面的文档假设用户是在一个 bash 终端（bash terminal）里工作的。

#### 2.1.1 安装系统环境

用 Synaptic Package Manager 或者 “sudo apt install” 安装下面的系统软件包：

- git
- gcc
- gfortran
- cmake
- python3-pip
- python-is-python3
- python3-venv
- openmpi-dev
- libxcb-cursor0
- libcfitsio-dev
- libgsl-dev

---

<sup>4</sup>DroPS 也提供“天图制作”的快速模拟器，但是该模拟器仅仅是一种近似，而非是针对某个具体实验的真实时序数据处理。

- libfftw3-dev
- libfftw3-mpi-dev
- libhealpix-dev

### 2.1.2 建立 Python 虚拟环境

在常用的工作目录（后面用 YourWorkPath 代替）下建立一个 Python 虚拟环境目录：

```
mkdir YourWorkPath/.work
```

（这里的虚拟环境目录名 .work 是可以随意选取的，但通常建议以 . 开头（默认为隐藏文件夹），以防止日常操作中误删除 Python 包）

然后创立一个 Python 虚拟环境：

```
python -m venv YourWorkPath/.work
```

激活虚拟环境：

```
source YourWorkPath/.work/bin/activate
```

在 Windows 系统里激活虚拟环境可能要在 cmd.exe 里执行

```
YourWorkPath/.work/Scripts/activate.bat
```

或者在 PowerShell 里执行

```
YourWorkPath/.work/Scripts/activate.ps1
```

当工作结束时，用

```
deactivate
```

来退出 Python 虚拟环境。

如果你没有其他的 Python 项目，建议把激活虚拟环境的命令放到 .bashrc 里

```
echo "source YourWorkPath/.work/bin/activate" >> ${HOME}/.bashrc
```

这样每次启动终端时会自动激活这个虚拟环境。

### 2.1.3 安装 python 依赖

在激活了上小节安装的 Python 虚拟环境的情况下，我们来安装 DroPS 依赖的 Python 包。

先更新 pip 到最新版以获得 Python 包的正确信息：

```
pip install --upgrade pip
```

进入到工作目录：

```
cd YourWorkPath
```

从 Github 上下载 DroPS 的源码：

```
git clone https://github.com/zqhuang/DroPS
```

然后进入 DroPS 目录

```
cd DroPS
```

用 pip 安装所有依赖包：

```
pip install -r requirements.txt
```

### 2.1.4 修改 pysm3 包

在上一节安装的依赖包里有一个叫 pysm3 的包，它的功能是模拟前景和 CMB 辐射。我们现在来修改它的源代码来提高模拟效率。虽然修改一个 Python 包的源码一般来说是不被推荐的操作，但是在效率和规范之间反复权衡之后，我们还是采取了这个“权宜之计”来获得更适合 DroPS 的高效模拟功能。在未来如果 pysm3 提供了更为灵活的功能，就可以避免这个不太符合 Python 理念的操作。另外，如果你只准备用 DroPS 的数据分析功能，而不打算使用 DroPS 的天图模拟功能，也可以跳过这小节的操作。

进入 DroPS 目录：

```
cd YourWorkPath/DroPS
```

把 pysm3 安装目录下的文件 models/cmb.py 移动并改名（这是为了备份以保证操作可逆）：

```
mv PATH_TO_pysm3/models/cmb.py cmb_backup.py
```

然后把 DroPS 目录下的 `cmb.py` 移动到 `pysm3` 安装目录的 `model` 文件夹中：

```
cp cmb.py PATH_TO_pysm3/models/
```

这里 `PATH_TO_pysm3` 代表 `pysm3` 的安装目录，在 Ubuntu 24.04.3 LTS 系统里，你可能会发现这个安装目录是

`YourWorkPath/.work/lib/python3.12/site-packages/pysm3`

在其他系统环境下，这个安装目录可能有所不同。在 linux 系统下，这个问题比较容易解决：可以先安装 `plocate`，

```
sudo apt install plocate
```

然后用

```
locate pysm3
```

就可以找到 `pysm3` 的安装路径。

## 2.2 参照数据组

在分析（无论是模拟的还是真实观测得到的）观测天图之前，必须模拟一批 CMB、噪声和前景的参照数据组。参照数据组的模拟代表了目前我们对现实（仪器、噪声统计性质、前景统计性质等）的最佳认知。参照数据组可以类比为机器学习中的训练样本<sup>5</sup>。但无论如何，需要注意参照数据组可能和真实数据之间存在系统性的差别。如何尽可能地避免这些可能的差别带来的影响，是 DroPS 一个重要的优化指标。

### 2.2.1 生成一个时序数据过滤的模拟器

处理地基 CMB 实验数据的关键步骤在于抑制时间序列数据中的地面和大气污染。要精确模拟这一滤波过程，需要详尽的巡天策略信息，而这在前期预测阶段往往难以获取。不过，我们对滤波产生的大致效应有充分的理解：该过程会显著抑制最终地图中的大尺度功率（球谐展开中的低阶极矩），并通过混合傅里叶模式引入非高斯关联。对于不需要精确模拟特定巡天策略的应用场景，DroPS 提供了一套“模拟滤波”工具来生成具有同样效应的滤波矩阵。注意这个方法是唯象近似，而非从原理上对真实实验的滤波做近似，所以只适合在实验的前期预测阶段对滤波效应的影响做大致的估算。另外，这个方法对计算资源的需求远低于真实滤波过程，所以在对模拟量有很大需求时也适合作为一种平衡计算成本和计算精度的一种选择。

运行

---

<sup>5</sup>注意这仅仅是一个类比，DroPS 并不使用深度学习等前沿机器学习技术。



```
python mock_filtering.py
```

来获得产生时序数据的滤波矩阵。后续只需要根据提示输入相应的描述滤波矩阵的参数：主要的两个参数是 Healpix 天图分辨率 (nside, 建议使用 128 做快速测试, 256 或者 512 做更为精确的测试) 和存储滤波矩阵的以 pickle 为后缀的文件名 (例如 filter\_128.pickle)。更精细的参数有  $\ell_{\text{cut}}$ ,  $f_{\text{low}}$ ,  $f_{\text{high}}$ ,  $\alpha_{\text{low}}$ ,  $\theta_{\ell}$ ,  $\theta_m$ 。它们的定义可以在3.1节找到。

### 2.2.2 生成参照数据组

有了滤波矩阵, 下一步就可以模拟生成参照数据组。不管是参照数据组还是观测数据的模拟, 都依赖于观测的策略和仪器精度, 因此必须针对某个观测实验进行模拟。DroPS 提供了四个“演示实验”: Test, AliCPT, SO, 和 CMBS4。其中 Test 是一个以快速测试为目的的纯假想实验, 仅包含四个频段, 以 nside=128 的低分辨率进行测试。其他三个演示实验 (AliCPT, SO, CMBS4) 的参数 (天区面积、频段分布、角分辨率、噪声水平等, 可以参考表 1) 都和同样名称的、已有或计划的实验 (AliCPT [5], SO-SAT [8] and CMB-S4-SAT [11]) 相近。在这个文档里我们主要以 AliCPT 这个演示实验为例子进行使用说明, 同时也会以更精确的 SO 这个演示实验为例来和文献中已经有的其他同类软件 [8] 做对比。不过要注意的时, 在写这个文档的时候, AliCPT 和 SO 这两个实验已经开始投入观测, 其实验参数和观测策略和 DroPS 中对应名称的演示实验不尽相同, 因此不能把 DroPS 的演示实验的测试结果当作真实实验的准确预测。

现在可以运行

```
python simulate.py AliCPT/AliCPT_sim_config.txt
```

来生成演示实验 AliCPT 的参照数据组。

参照表1来阅读模拟的设置文件 AliCPT/AliCPT\_sim\_config.txt, 理解参数时如何在设置文件里被定义的。如果在运行 simulate.py 时没有从命令行传入任何其他参数, 它仅会产生参照数据组——也就是基于噪声模型和宇宙学分别模拟的大量用不同随机种子生成的噪声和 CMB 天图, 以及基于前景模型产生的一张前景天图。DroPS 默认会使用两个不同的  $r$  值 ( $r = 0$  以及  $r = r_1 = 0.03$ , 其中  $r_1$  的值允许在设置文件里修改) 来生成两组不同的 CMB 天图。这是为了避免因为对时序数据滤波的建模不准确性带来的系统误差。详细的原因可以参考3.3。

在生成参照数据组时, DroPS 默认用最简单的 ['d0', 's0'] 模型<sup>6</sup>给出一张固定的前景天图。这反映了目前测量原初引力波所面临的最重要的技术瓶颈: 我们既没有一个从第一性原理出发的理论模型、也没有足够的多波段测量数据来建立一个可靠的描述前景

<sup>6</sup>['d0', 's0'] 模型假设固定的尘埃热辐射谱指数  $\beta_d = 1.54$  以及同步辐射谱指数  $\beta_s = -3$ ; 谱指数的定义见3.4节。

Table 1: DroPS 演示实验的关键参数，其中  $\ell_{\text{knee}}$  and  $\alpha_{\text{knee}}$  在3.2节里可以找到定义

	中心频率 (GHz)	分辨率 (FWHM) (arcmin)	噪声 $\sigma_{\text{white}}$ ( $\mu\text{K-arcmin}$ )	$\ell_{\text{knee}}$	$\alpha_{\text{knee}}$
Test ( $f_{\text{sky}} \approx 0.1$ )	30	60	1	60	-1.7
	95	20	1	60	-1.7
	150	20	1	60	-3
	270	10	1	60	-3
AliCPT ( $f_{\text{sky}} \approx 0.14$ )  (+Planck)	27	97	47	30	-2.4
	40	65	34	30	-2.4
	90	16.2	2.2	50	-2.5
	150	9.7	3.3	50	-3
	217	4.9	50	0 <sup>+</sup>	-3
SO ( $f_{\text{sky}} \approx 0.1$ )	27	91	33	15	-2.4
	39	63	22	15	-2.4
	93	30	2.5	25	-2.5
	145	17	2.8	25	-3
	225	11	5.5	35	-3
	280	9	14	40	-3
CMBS4 ( $f_{\text{sky}} \approx 0.1$ )	30	72.8	2.5	60	-1.7
	40	72.8	3.15	60	-1.7
	85	25.5	0.622	60	-1.7
	95	22.7	0.552	60	-1.7
	145	25.5	0.87	60	-3
	155	22.7	0.948	60	-3
	220	13	2.46	60	-3
	270	13	4.22	60	-3

能谱涨落的统计模型。一个衡量 DroPS 以及类似软件的重要指标就是,  $r$  的测量结果不能严重依赖于真实前景和参照数据组中采用的前景模型的差别。对实际观测数据而言, 我们无法不依赖于模型地得知天图中包含的“真实前景”。因此对 DroPS 的性能的测试要依靠前景模型已知 (但可以参照数据采用的  $['d0', 's0']$  不同) 的模拟数据。

下一节我们要展示如何用模拟数据测试 DroPS 的性能。在此之前, 用户可以进一步参考 pysm3 的官方说明文档 <https://pysm3.readthedocs.io/> 来熟悉  $['d0', 's0']$  以及其他前景模型的定义。

## 2.3 分析天图

在上一节我们按照噪声模型、宇宙学模型 (包括两个不同的  $r$  值) 和  $['d0', 's0']$  前景模型模拟生成了参照数据组。现在我们要用同一个噪声模型、一个用户假定的宇宙学 (用户假定的  $r$  值是任意的, 和参照数据组采用的两个  $r$  值完全无关)、以及用户指定的前景模型 (可以和参照数据组采用的  $['d0', 's0']$  不同) 来模拟“观测天图”。DroPS 的主要功能是在不知道用户输入的宇宙学模型和前景模型的情况下, 把“观测天图”和参照数据组做对比, 并推断出用户假定的  $r$  的值。从概念上可以把参照数据组类比为机器学习中的训练样本, 而把“观测天图”类比为测试样本。在涉及机器学习的研究中, 往往采用大量的测试样本来评估机器学习的准确率。同样地, 我们可以用不同的随机种子、宇宙学模型、前景模型来生成多张“观测天图”, 来评估 DroPS 测量的  $r$  值的系统偏差的大小。

### 2.3.1 单张模拟观测天图的分析

我们先来生成一张“观测天图”:

```
python simulate.py AliCPT/AliCPT_sim_config.txt maps/AliCPT_ 0.01 999
```

现在我们从命令行传递给 simulate.py 多个参数。第一个参数仍然是设置文件。第二个参数 maps/AliCPT\_ 是存储观测天图时采用的文件名前缀 (例如 40GHz 的天图将被保存为 AliCPT/maps/AliCPT\_40GHz.npy)。第三个参数 0.01 是用户假设的  $r$  值。第四个参数 999 是随机种子。如果用户想采用和设置文件中的  $['d0', 's0']$  不同的前景模型, 有两种方案: 一种是直接修改设置文件, 把  $['d0', 's0']$  替换成例如  $['d1', 's1']$ 。另外一种方案是在命令行添加第五个参数, 例如  $d1s1$ 。

现在用

```
python mainpipe.py AliCPT/AliCPT_ana_config.txt maps/AliCPT_
```

来分析我们刚刚生成的“观测天图”。我们传递给 mainpipe.py 的第一个参数是设置文件, 注意它并不包含噪声水平以及用户用于生成“观测天图”所采用的  $r$  值和前景模型

这些信息。DroPS 的任务就是利用参照数据组来获得噪声的统计性质，并通过对比参照数据组和“观测天图”来推断用户采用的  $r$  值。

在这一步，用户可以简单地阅读设置文件 AliCPT/AliCPT\_ana\_config.txt 来获得对 DroPS 分析天图时涉及的设置参数的大致了解。

当计算结束时，张标比  $r$  以及部分其他宇宙学和前景参数的后验统计将会在屏幕上显示出来。更多的结果会被记录到 AliCPT/results。这个保存结果的路径是在设置文件 AliCPT/AliCPT\_ana\_config.txt 里指定的。

### 2.3.2 多张模拟观测天图的分析

受宇宙学方差的影响，对单张模拟观测天图的分析给出的  $r$  的后验分布的平均值并不会精确地等于用户输入的  $r$  值，具体的差异依赖于随机种子和 DroPS 本身具有的系统误差。要把（我们关心的）DroPS 的系统误差和（我们不关心的）宇宙学方差分离开来，就要用不同的随机种子多次重复这种分析，并把大量重建出来的  $r$  的后验分布平均值的平均（即平均的平均）和用户输入的  $r$  做比较。

我们打算采用  $r = 0.01$ ，前景模型 [“d1”, “s1”] 并让随机种子遍历 1 到 100。在命令行手动输入 100 次指令显然不是明智的做法，因此我们用一个 DroPS 自带的 bash 脚本 test\_bias.sh 来代劳。直接运行

```
./test_bias.sh
```

即可完成任务。如果存在运行权限问题，就先用

```
chmod 700 test_bias.sh
```

修改权限。

下面给出 test\_bias.sh 的具体内容

Listing 1: content of test\_bias.sh

```
for i in `seq 100`
do
    python simulate.py AliCPT/AliCPT_sim_config.txt maps/s${i}_ 0.01 ${i} d1s1
    python mainpipe.py AliCPT/AliCPT_ana_config.txt maps/s${i}_ logfile.txt
done
```

这里的 bash 变量  $\{i\}$  遍历了 1 到 100。我们额外传递给 simulate.py 一个命令行参数 'd1s1' 来指定前景模型；额外传递给 mainpipe.py 一个命令行参数 logfile.txt 来指定存储结果（每次模拟得到的  $r$  的后验分布平均和标准差）的文件。

当运行 bash 脚本结束（这在个人电脑上可能要几个小时）后，可以把 logfile.txt 中存储的结果用下述命令画出来

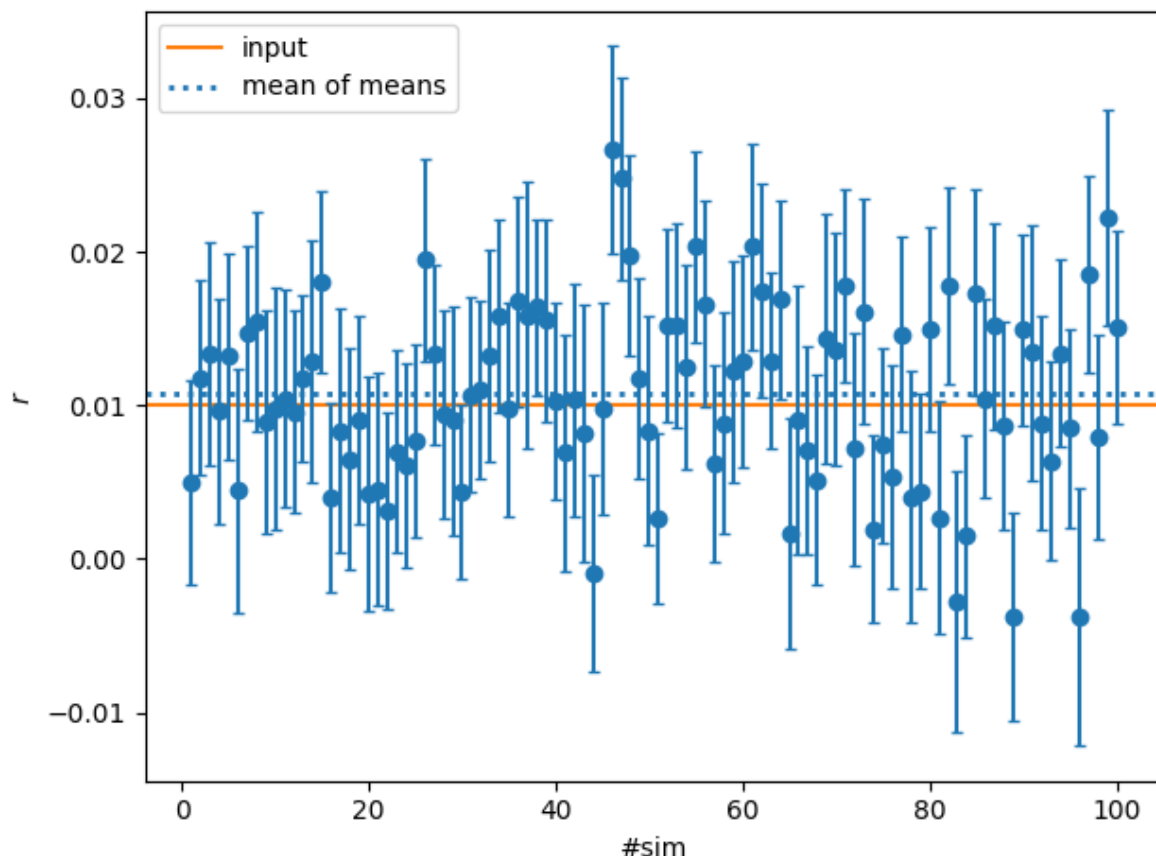


Figure 1: 100 张不同随机种子产生的模拟观测天图的后验分布（均值和标准差）

```
python utils/plot_rs.py logfile.txt 0.01
```

这里我们传递给了画图脚本用户采用的  $r = 0.01$ ，是为了把它和 logfile.txt 中存储的所有  $r$  的后验分布平均的平均值做比较。

如果 DroPS 推断的  $r$  的值是几乎无偏的（系统偏差远小于统计误差），那么 logfile.txt 中存储的所有  $r$  的后验分布平均的平均值（mean of means）将和用户采用的  $r$  值非常接近。这在图1中得到体现。

### 2.3.3 真实观测数据上的应用

每一个实验都有自己的天图制作和噪声模拟数据管线，可以提供噪声和 CMB 的参照数据组；另外，实验也会提供观测天图。只要把这些文件按照 DroPS 的命名规则进行命名，就可以用 DroPS 进行数据分析。

在 DroPS 中默认是有  $r = 0$  和  $r = r_1 = 0.03$  的两组 CMB 参照数据。如果实验仅提供  $r = 0$  的 CMB 参照数据组，要在分析的设置文件中把 cosmo\_r1 取值为 None。不过，在计算资源允许的情况下，最理想的处理方式还是再生成一组  $r = r_1$  的 CMB 参照数据组。

真实实验的每个频段都有一定的带宽，即有一个频率分布函数  $f(\nu)$ 。在这种情况下，可以在设置文件中定义一个文件夹的路径 `band_weights_root`，并把每个频段的分布函数存储在这个文件夹下。这些文件的命名和存储格式，可以参考 `Test` 这个演示实验。

### 2.3.4 和文献中同类软件比较

DroPS 曾被用于 AliCPT 实验的数据挑战并获得了和模拟输入一致的结果 [6]。

这里我们进一步把 DroPS 和文献中其他同类软件进行比较。在这类测试中，比较有代表性的是对 SO-SAT 的模拟数据的分析和比较 (Wolz et al. 2024 [8])。DroPS 里自带的演示实验 SO 对应于 SO-SAT 在乐观假设下的预测。在 Wolz et al. 2024 中，对比了三组不同的数据分析软件对基于  $r = 0$  的模拟数据的分析。在分析中，允许  $r < 0$  的参数范围（在  $r < 0$  时让理论预言的原初 B 模功率谱反号），DroPS 也允许进行这样的分析<sup>7</sup>。

图2展示了对 SO 的分析结果。从上至下，分别是对基于 `['d0', 's0']`, `['d1', 's1']`, and `['dm', 'sm']` 三种前景模型各 250 张模拟观测天图的分析结果<sup>8</sup>。左侧展示了  $r$  的后验平均的直方图，以及这些  $r$  后验平均的平均值 (mean of means)；这个平均值和模拟输入  $r = 0$  的差异体现的是 DroPS 分析方法的系统误差。右侧展示了  $r$  的后验标准差的直方图。可以看到，在所有情况下 DroPS 分析方法的系统误差都远小于统计误差，综合表现要稍优于 Wolz et al. 2024 中的三种分析方法 (pipeline A, pipeline B 和 pipeline C)。DroPS 得到的标准差的分布也和 Wolz et al. 2024 中的三种分析方法得到的结果相洽。

## 2.4 天图成分分离

在这一节我们考虑天图的成分分离。目标是从多频段的观测天图中获得 CMB 的 E 模和 B 模天图。

### 2.4.1 基于似然函数的天图成分分离方法

最常见的成分分离方法可能是内部线性组合 (Internal Linear Combination, ILC) 算法及其各种变体。其基本思想是，通过对多频段天图进行线性组合，来分离出一个频率依赖性已知的信号。若要在像素空间中进行此操作，必须将各频段的天图平滑至一个共同的分辨率。

然而，对于地基 CMB 实验，ILC 或类 ILC 方法面临的一个关键挑战在于：时序数据滤波与波束卷积是不可对易的操作。因此，将各频段的天图平滑至共同分辨率，会不可避免地改变它们的光谱能量分布。这一固有的局限性解释了，尽管在忽略时序数据滤

<sup>7</sup>只要在设置文件中取 `r_lowerbound` 为一个足够小的负值就行。

<sup>8</sup>这里的 `['dm', 'sm']` 并不是 `pysm3` 的自带模型，需要自己按照 Wolz et al. 2024 [8] 的说明生成。

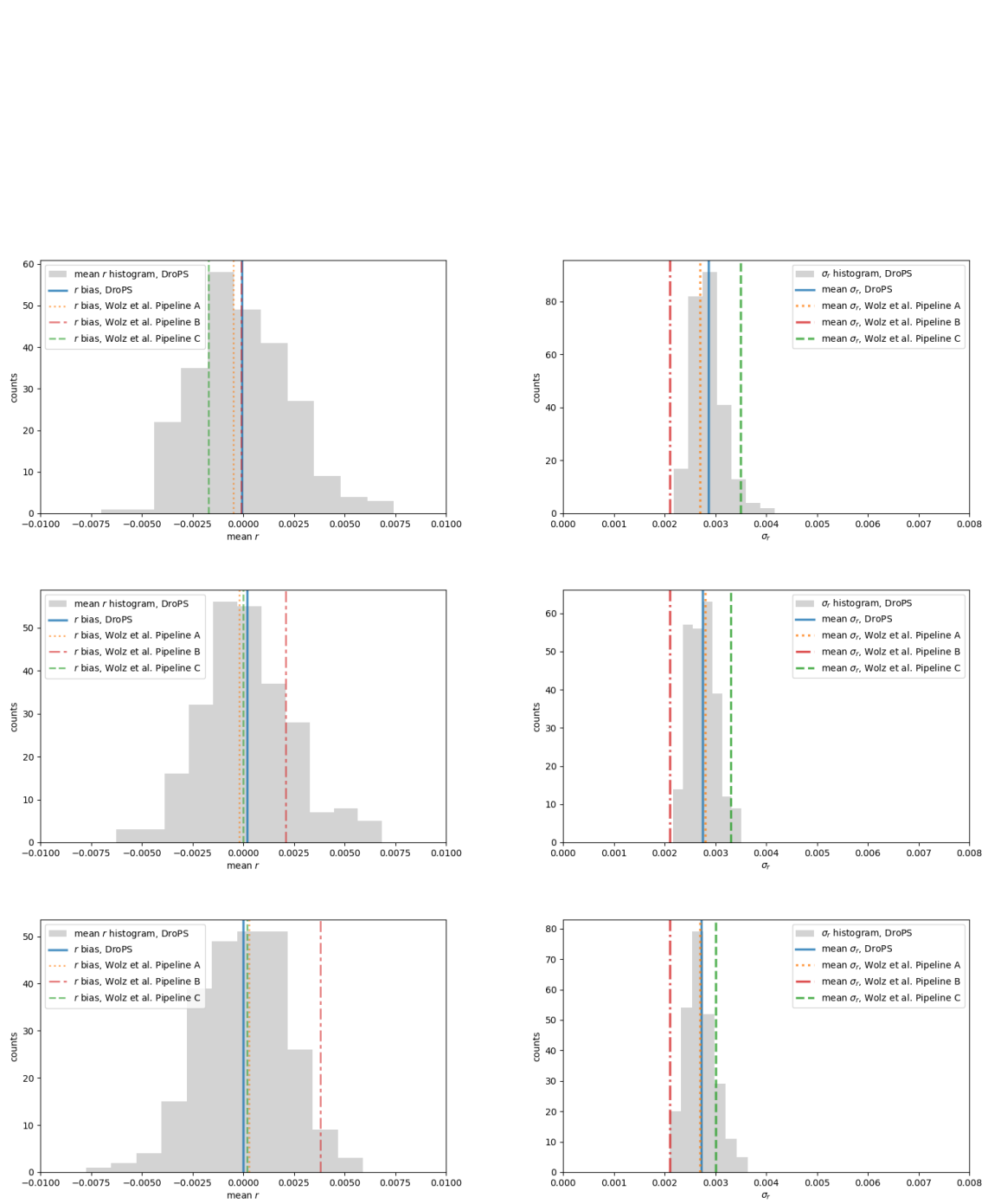


Figure 2: DroPS 和 Wolf et al. 2024 [8] 中的三种分析方法的比较：从上至下，分别是对基于 ['d0', 's0'], ['d1', 's1'], and ['dm', 'sm'] 三种前景模型的各 250 张模拟观测天图的分析结果。左侧是  $r$  的后验平均的直方图；右侧是  $r$  的后验标准差的直方图。



波效应的模拟研究中 ILC 方法非常流行 [8], 但在真实数据分析中, 例如 BICEP/Keck 合作组所进行的分析 [12], 它通常会被计算量更大、基于似然分析的方法所取代。

DroPS 的 `compsep.py` 脚本提供了天图的成分分离功能。为了展示它如何工作, 我们把前面生成的 AliCPT 参照数据组的零号模拟的 CMB、噪声和前景天图相加来生成一张“观测天图”。这样做的好处是我们知道输入的 CMB 天图 (包括未经过时序数据滤波的原始输入天图) 是什么, 在完成天图成分分离后有一个“正确答案”可以参考。下面的脚本直接提供了从参照数据组生成观测天图的功能:

```
python utils/combine_sim.py AliCPT/AliCPT_sim_config.txt maps/r3AliCPT0_ 0
```

从命令行传递给脚本的第一个参数是模拟的设置文件 (包含了所有关于参照数据组的信息); 第二个参数 `maps/r3AliCPT0_` 是用户指定的生成天图的命名前缀 (不要和之前生成过的天图前缀重复); 第三个参数 `0` 是指使用参照数据组 `0` 号模拟。

有了以 `maps/r3AliCPT0_` 为前缀的模拟天图之后, 我们用 `compsep.py` 来对它进行成分分离:

```
python compsep.py AliCPT/AliCPT_sim_config.txt maps/r3AliCPT0_
```

要注意这个步骤需要大约 100GB 的存储空间和 10GB 的内存。

该代码采用随机梯度朗之万动力学方法, 用于搜寻第 3.6 节所述似然函数的最大值。在个人计算机上首次运行 `compsep.py` 通常需要数天时间。当梯度模板构建完成后, 代码运行速度将显著提升, 在个人计算机上通常仅需数小时即可完成。一般而言, 需要多次运行该程序才能获得充分收敛的 CMB  $B$  模天图。

图 3 展示了运行结果。其中  $E$  模天图的还原效果非常理想, 而  $B$  模天图中仍存在显著的残差噪声。另外, 时序数据的过滤也使得最大角尺度信息有一定损失<sup>9</sup>。

在真实数据分析中, 梯度模板的计算涉及天图制作流程, 因此构建所需时间更为漫长 (无法用个人电脑完成)。值得庆幸的是, 梯度模板可以通过并行方式进行构建。通过执行以下指令即可构建特定多极矩范围  $[\ell_{\min}, \ell_{\max}]$  内的模板:

```
python compsep.py AliCPT/AliCPT_sim_config.txt maps/r3AliCPT0_ lmin lmax
```

借助现代架构普遍支持的并行 I/O 技术, 该任务可被分解为多个具有不同 `lmin` 与 `lmax` 取值的子任务。

<sup>9</sup>CMB 的  $E$  模式在最大角尺度上的功率谱比较低, 所以这个效应在 CMB 的  $E$  模式天图中相对不那么明显



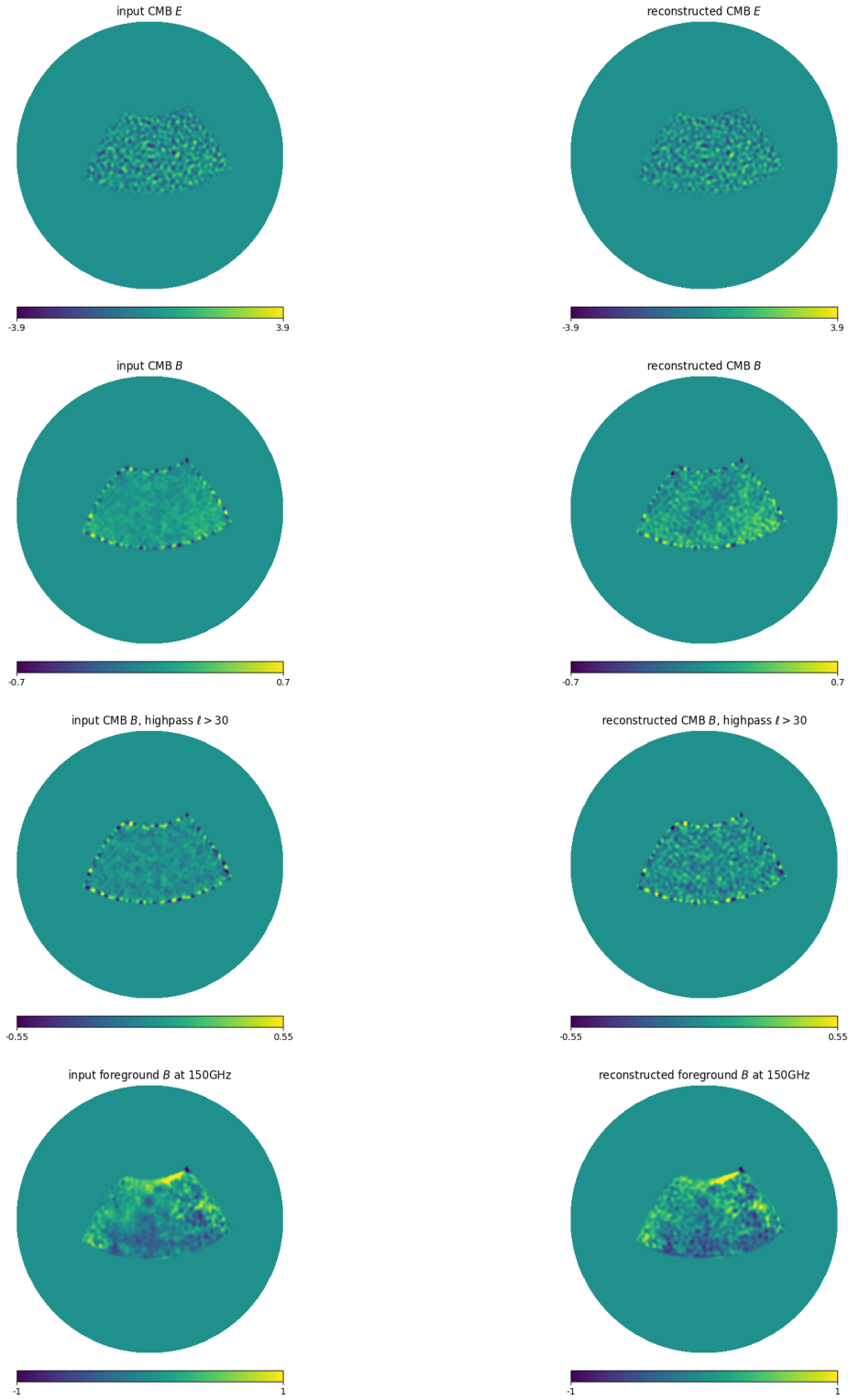


Figure 3: AliCPT 模拟观测天图的成分分离结果；为了聚焦于我们感兴趣的大角尺度信息，所有天图都经过光滑（取  $\ell \leq 160$  的截断）

### 2.4.2 可视化

对于仅观测部分天区的地面 CMB 实验，为节省磁盘空间，DroPS 仅保存观测天区内的像素。这类以 npy 格式存储的天图无法直接通过标准 Healpix 工具进行可视化。DroPS 提供了专用工具用于显示：

脚本 `utils/viewfits.py` 可显示 fits 格式天图：

```
python utils/viewfits.py NHmask_G_128.fits
```

对于包含斯托克斯参量 I、Q、U 的天图文件 `example_IQU.fits`，可通过以下命令显示指定分量（`mapind=0,1,2` 分别对应 I、Q、U 参量）：

```
python utils/viewfits.py example_IQU.fits mapind
```

若要显示 npy 格式的掩膜天图（需配合掩膜文件 `mask.fits`），使用 `utils/viewnpy.py`：

```
python utils/viewnpy.py mask.fits example_IQU.npy mapind
```

显示 E/B 模式分量需使用 `utils/viewEB.py`：

```
python utils/viewEB.py mask.fits example_IQU.npy E/B
```

### 3 软件技术细节

在介绍软件的技术细节的这一节，我们假设用户具有一定的宇宙学背景知识，并熟悉常用 Healpix [13] 等软件。

DroPS 通过将多极矩范围划分为若干的  $\ell$  区间，并利用 NaMaster [14] 计算掩膜天图的带功率<sup>10</sup>，以此分析功率谱。在测量张标比  $r$  时，分析通常聚焦于  $20 \lesssim \ell \lesssim 200$  的范围。此时  $n_{\text{side}}=256$  的天图分辨率已足够满足需求，甚至  $n_{\text{side}}=128$  的分辨率也可用于快速估算。但需要特别注意，在执行诸如改变天图分辨率（Healpix 的 `ud_grade`）或旋转至不同坐标系等操作时，这些处理可能在低分辨率天图中引发显著的 E-B 泄漏。对于涉及此类操作的分析，我们建议采用  $n_{\text{side}}=512$  或者更高分辨率。

DroPS 用于评估  $r$  以及其他宇宙学和前景参数的似然分析所用的数据向量由滤波天图得到的带功率构成。其构建方式如下：不同频段间的带功率直接作为交叉功率谱计算，而对单一频段，则通过该频段内不同观测季天图两两间的交叉功率获得。这种设计确保了噪声带功率的期望值在所有情况下均为零，从而有效消除了不完善噪声建模带来的偏差。

假设有  $n_{\text{bin}}$  个  $\ell$  区间， $n_\nu$  个频段，以及  $n_{\text{fld}}$  种关联场对（比如温度自关联 TT，温度和 E 极化关联 TE，E 极化自关联 EE，B 极化的自关联 BB 等）。DroPS 所用的数据矢量  $D_\ell^{XY}(\nu_1, \nu_2)$  依次按下列指标排序：

- $n_{\text{bin}}$  个  $\ell$  区间的中心值；
- $n_\nu(n_\nu + 1)/2$  个有序频率对  $(\nu_1, \nu_2)$  ( $\nu_1 \leq \nu_2$ )；
- $n_{\text{fld}}$  关联场对  $XY$ 。

如前所述，对  $\nu_1 = \nu_2$ ，带功率  $D_\ell^{XY}(\nu_1, \nu_2)$  是通过对不同观测季之间的天图之间互关联计算平均得到的。

由此可得完整数据向量的总长度为  $N_d = \frac{1}{2}n_{\text{fld}}n_{\text{bin}}n_\nu(n_\nu + 1)$ 。针对计算或存储资源受限的  $r$  测量场景，我们建议仅使用  $BB$  谱（即取  $n_{\text{fld}} = 1$ ），因为该谱包含近乎全部的  $r$  信息。

DroPS 评估参数的基本方法是通过评估观测数据向量  $D_\ell^{XY}(\nu_1, \nu_2)$ （抽象记为  $D^{\text{obs}}$ ）在给定模型预测（抽象记为  $D^{\text{model}}$ ）条件下的似然度。在高斯近似下，似然函数为：

$$\mathcal{L} = \frac{1}{(2\pi)^{N_d/2} \sqrt{\det \text{Cov}}} \exp \left[ -\frac{1}{2} (\mathbf{D}^{\text{obs}} - \mathbf{D}^{\text{model}})^T \text{Cov}^{-1} (\mathbf{D}^{\text{obs}} - \mathbf{D}^{\text{model}}) \right], \quad (3)$$

模型数据向量又可以拆解为：

$$\mathbf{D}^{\text{model}} = \mathbf{D}^{\text{CMB}} + \mathbf{D}^{\text{fg}}, \quad (4)$$

其中  $D^{\text{CMB}}$  and  $D^{\text{fg}}$  分别表示来自 CMB 和前景的带功率期望值。由于完全采用交叉关联计算，噪声带功率的期望值为零，所以噪声对  $D^{\text{model}}$  没有贡献。总协方差矩阵  $\text{Cov}$  包含来自噪声、前景和 CMB 的所有方差成分。

<sup>10</sup>带功率表示给定  $\ell$  区间内功率谱的平均值

在介绍噪声、CMB 和前景模型以及估计关联矩阵 Cov 的方法之前，我们先介绍与之密切相关的天图制作流程和时序数据的过滤。

### 3.1 天图制作和时序数据过滤

地面观测数据会受到大气辐射和仪器  $1/f$  噪声等低频干扰的严重污染。地基 CMB 实验中的一个关键预处理步骤是时序数据滤波，该技术旨在从主导性的、结构化的地面噪声中分离出微弱的 CMB 信号。时序数据滤波在时域中运行，通过例如三阶多项式拟合等方法扣除时序中的低频成分。该方法在抑制大尺度噪声的同时，力求尽可能保留在小尺度上的 CMB 信号。

虽然这是一个线性过程，但时序数据滤波等效于施加一个稠密矩阵运算，若缺乏实验扫描策略的详细信息则难以建模。因此，DroPS 不包含针对特定实验的完整天图制作流程，而是提供简化的“天图制作”模拟器。这些模拟器通过结合像素空间和谐空间中的若干线性操作，根据巡天掩膜和一组描述时序数据滤波核心特征的参数，实现对时序数据滤波的主要物理效应的近似模拟。

具体而言，模拟器通过下列对输入天图的一系列线性操作来模拟时序数据滤波的主要物理效应：

1. 在谐波空间生成一个过滤矩阵。这个矩阵依赖于用户指定的 6 个参数： $\ell_{\text{cut}}$ ,  $f_{\text{low}}$ ,  $f_{\text{high}}$ ,  $\alpha$ ,  $\theta_\ell$  和  $\theta_m$ 。具体地说，它是

$$M_{\ell'm',\ell m} = \begin{cases} \sqrt{[1 - 2(\theta_\ell + \theta_m)] F_\ell} & \text{如果 } \ell' = \ell, m' = m; \\ \sqrt{F_\ell \theta_\ell} G_{\text{random}} & \text{如果 } |\ell' - \ell| = 1, m' = m; \\ \sqrt{F_\ell \theta_m} G_{\text{random}} & \text{如果 } \ell' = \ell, |m' - m| = 1; \\ 0, & \text{其他情况} \end{cases}$$

这里的

$$F_\ell \equiv f_{\text{high}} \left\{ 1 - f_{\text{low}} \exp \left[ - \left( \frac{\ell}{\ell_{\text{cut}}} \right)^\alpha \right] \right\},$$

$G_{\text{random}}$  是一个（服从标准正态分布的）高斯随机数。每个矩阵元里的  $G_{\text{random}}$  都独立采样。

2. 用 2 度大小的 C2-类卷积核把天区掩膜（mask）的边缘光滑化 [15]。
3. 把输入天图和边缘光滑后的掩膜逐像素相乘。
4. 把相乘后的图进行球谐函数展开得到展开系数  $a_{\ell m}$ 。
5. 用第 1 步中生成的过滤矩阵和球谐展开系数组成的矢量相乘：

$$\tilde{a}_{\ell'm'} = \sum_{\ell,m} M_{\ell'm',\ell m} a_{\ell m}.$$

6. 把  $\tilde{a}_{\ell'm'}$  球谐逆变换回像素空间。
7. 把像素空间的图用原始（未经过边缘光滑的）掩膜覆盖。

当运行

```
python mock_filtering
```

后，会得到一个 pickle 文件来保存步骤 1 中的过滤矩阵。

当模拟参照数据组或者“观测天图”时，simulate.py 会从设置文件里读取保存过滤矩阵的 pickle 文件名。然后读取该文件里的过滤矩阵。注意对一个固定的演示实验，过滤矩阵和天空掩膜是固定的。也就是说，步骤 1 和 2 仅需要执行一次，真正需要重复对每个天图执行的是步骤 3-7。

上面描述的天图制作模拟器不涉及对真正时序数据的处理，所以对计算资源的需求远低于真实实验中的天图制作管线。DroPS 提供这样的模拟器的目的是唯象地刻画时序数据滤波的主要物理效应（大尺度功率谱的压低和不同球谐模之间的混合），而并不试图精确重现任何具体实验的天图制作管线。

在数据分析阶段，DroPS 并不要求提供天图制作管线中的滤波矩阵（在真实实验中这是比较困难的），而是通过一个简化的作用于带功率的耦合矩阵  $F_{\ell\ell'}$  来对滤波净效应进行唯象建模。该矩阵的主对角线元素描述各  $\ell$  区间内的功率抑制效应，次对角线元素捕捉相邻区间之间的耦合效应，其余矩阵元均假设为零。在默认设置下，DroPS 通过对比基础模拟数据中经过滤波与未经滤波的噪声天图的带功率来对  $F_{\ell\ell'}$  进行最小二乘拟合。

需要重点说明的是：虽然时序数据滤波在天图层面是线性操作，但其对功率谱产生的影响并非严格线性。因此，作用于带功率的耦合矩阵  $F_{\ell\ell'}$  实质上是一种近似模型。基于此，DroPS 在算法设计中刻意降低了对  $F_{\ell\ell'}$  的依赖程度，以确保分析的稳健性。（更多细节参见第 3.3 节）。

## 3.2 噪声模型

DroPS 采用文献中常用的噪声模型 [16, 8]，其温度噪声为

$$N_\ell = N_{\text{white}} \left[ 1 + \left( \frac{\ell}{\ell_{\text{knee}}} \right)^{\alpha_{\text{knee}}} \right]. \quad (5)$$

这里的  $1/f$  噪声成分由参数  $\ell_{\text{knee}}$  和  $\alpha_{\text{knee}}$  描述。白噪音水平  $N_{\text{white}}$  和设置文件里给出的  $\sigma_{\text{white}}$  (见表 1) 由下式联系：

$$N_{\text{white}} = \sigma_{\text{white}}^2, \quad (6)$$

但是要注意上式中的  $\sigma_{\text{white}}$  的单位必须转换为  $\mu\text{K-radian}$ <sup>11</sup>。

<sup>11</sup> 单位转换公式为  $1 \mu\text{K-arcmin} = \frac{\pi}{10800} \mu\text{K-radian}$ 。

Table 2: DroPS 模拟参照数据组时默认使用的宇宙学参数

参数	定义	默认值
$\Omega_b h^2$	重子密度	0.02242
$\Omega_c h^2$	冷暗物质密度	0.11933
$\theta_{\text{MC}}$	声视界在最后散射面上的张角	0.0104101
$\tau_{\text{re}}$	再电离光深	0.0561
$\ln(10^{10} A_s)$	原初标量扰动大小的对数	3.047
$n_s$	原初标量扰动的谱指数	0.9665
$r$	张标比	$r_0 = 0$ 和 $r_1 = 0.03$

DroPS 自带的演示实验中，默认极化噪声功率谱是温度噪声功率谱的 2 倍，也就是  $N_\ell^{EE} = N_\ell^{BB} = 2N_\ell^{TT}$ 。虽然这对于实际实验未必是完全准确的，不过在大多数情况下近似成立。

注意 DroPS 仅在生成模拟数据时使用(5)式给出的噪声模型。在数据分析阶段，DroPS 并不对噪声模型做任何模型假设，而是直接从参照数据组中计算噪声带功率的关联矩阵。噪声带功率的理论期待总是为零，因为 DroPS 只使用不同频段之间的，或者同一频段内不同观测季的天图之间的交叉关联。

### 3.3 CMB 模型

#### 3.3.1 宇宙学

在做宇宙学分析时，DroPS 假设标准的 Lambda 冷暗物质 ( $\Lambda$ CDM) 模型，但是允许张标比  $r$  非零。表2给出了演示实验中默认使用的宇宙学参数。

和可以从参照数据组中直接提取的噪声统计不同，CMB 的统计依赖于宇宙学参数的真值（而不是参照数据组中采用的默认值），特别是  $r$  的真值。这个本质上的差别，再加上 CMB 透镜效应，导致了 CMB 对带功率关联矩阵的贡献以及 CMB 带功率的理论模型比噪声要复杂很多。我们先来介绍 CMB 透镜效应的处理方法。

#### 3.3.2 去透镜化

大尺度结构的引力透镜效应会使原始宇宙微波背景辐射发生畸变，平滑其声学振荡峰，并将部分 E 模式偏振转化为次级 B 模式信号。这种透镜效应产生的 B 模式在小尺度上成为主导前景，掩盖了来自暴胀引力波的更微弱的原初 B 模式。虽然透镜 B 模式功率的期望值可以通过理论预测并予以扣除，但我们观测宇宙中的实际透镜势会围绕该均值随机涨落（宇宙学方差的体现）。对于以  $\sigma_r \lesssim 10^{-3}$  为测量精度的目标而言，这种宇宙学方差造成的、不可预测的涨落将成为不可忽略的污染源。

去透镜技术通过从高分辨率 CMB 数据或外部示踪物中统计重构实际透镜畸变的引力势，并将其产生的光线偏折效应尽可能地从观测天图中移除，从而抑制干扰性的透镜



B 模式功率，显著增强对原初信号的探测灵敏度。

在 DroPS 中，其实现方式假设已有通过外部资源获取的理想透镜势天图。尽管原则上此类天图可实现完全去透镜，但实际限制使得完美去除无法实现。为此，DroPS 引入了去透镜效率参数  $\epsilon_{\text{delens}}$ （配置文件中为 `delens_fac`），用以指定需去除的透镜效应比例。具体而言，该过程通过在实施去透镜操作前，将输入透镜势天图按  $(1 - \epsilon_{\text{delens}})$  比例缩放来实现。

### 3.3.3 宇宙学参数限制

DroPS 默认除了张标比  $r$  之外的宇宙学参数都被其他实验测量得非常准确，从而在限制模型参数时可以让它们固定不动。用户可以在参数文件中设置 `vary_cosmology` 为 `True` 来让所有的宇宙学参数变动。不过，一般而言地基小口径望远镜实验的数据无法很好地限制所有宇宙学参数，所以为了避免出现严重的参数简并，DroPS 用一个 Planck+BAO 的关联矩阵 [1] 作为先验假设来保证参数都在目前实验允许的范围内跑动<sup>12</sup>。

DroPS 的 CMB 参照数据组在  $r = 0$  和  $r = r_1 > 0$  两种情况下分别生成。通过对参照数据组计算平均，我们可以在这两种情况下计算 CMB 带功率的平均，也就是  $\langle D_\ell^{\text{CMB}}(r = 0) \rangle$  和  $\langle D_\ell^{\text{CMB}}(r = r_1) \rangle$ 。DroPS 推荐的  $r_1$  默认值是 0.03 (大约是目前 90% 置信度上限)，不过用户可以在设置文件里改变它。

在标准模型里，假设除了  $r$  之外的宇宙学参数是固定的 ( $\theta = \theta_{\text{fid}}$ )，CMB 的带功率就近似地线性依赖于  $r$ 。因此，DroPS 用线性插值的方法来获得 CMB 带功率的主要部分，然后对剩余部分进行修正。具体就是：

$$D_\ell^{\text{CMB}} = (1 - w) \langle D_\ell^{\text{CMB}}(r = 0) \rangle + w \langle D_\ell^{\text{CMB}}(r = r_1) \rangle + \Delta_{\text{correction}}. \quad (7)$$

这个插值权重  $w$  定义为  $w \equiv \max \left[ \min \left( \frac{r}{r_1}, 1 \right), -1 \right]$ ，也就是在  $|r| > r_1$  时对插值权重做了截断（不做外插）。因为  $\langle D_\ell^{\text{CMB}}(r = 0) \rangle$  和  $\langle D_\ell^{\text{CMB}}(r = r_1) \rangle$  已经包含了 E 到 B 泄漏等效应，这个插值项就自动包含了所有非宇宙学效应的主要贡献。

修正项  $\Delta_{\text{correction}}$  体现了其他宇宙学参数的影响，或者  $|r| > r_1$  情况下的修正。它可以写作：

$$\Delta_{\text{correction}} = \sum_{\ell'} F_{\ell\ell'} \left[ D_{\ell'}^{\text{CMB}}(r, \theta) - (1 - w) D_{\ell'}^{\text{CMB}}(r = 0, \theta_{\text{fid}}) - w D_{\ell'}^{\text{CMB}}(r = r_1, \theta_{\text{fid}}) \right]. \quad (8)$$

注意上式中的 CMB 带功率都是用 CAMB 计算得到的理论值 [17]。

在最为可能的情况 ( $0 < r < r_1$ ，并且其他宇宙学参数被精确测量) 下，上述算法中的修正项比较小。从而尽可能地避免了因时序数据过滤建模 ( $F_{\ell\ell'}$  矩阵) 不准确性带来额外的系统误差。

<sup>12</sup> 用户可以修改 `mainpipe.py` 来取消这个先验假设，但是一般不推荐这样做。

### 3.4 前景模型

#### 3.4.1 尘埃热辐射和同步辐射的基本模型

宇宙微波背景偏振测量中的主要弥漫前景——银河系热尘埃辐射与同步辐射——具有与 CMB 黑体辐射谱截然不同的能谱分布。该光谱差异构成了成分分离方法的基础。显然，至少需要三个频段的观测数据才能有效分离三种成分（CMB、尘埃辐射与同步辐射）的信号。

银河系尘埃热辐射近似服从修改黑体辐射（modified black-body, MBB）谱。以 CMB 背景温度为参照黑体辐射谱来描述等效的温度扰动，MBB 谱可以写作：

$$W_d(\nu; \nu_{\text{ref}}, \beta_d) = \left( \frac{\nu}{\nu_{\text{ref}}} \right)^{\beta_d-1} e^{\frac{h(\nu_{\text{ref}}-\nu)}{k_B T_{\text{CMB}}}} \left( \frac{e^{\frac{h\nu}{k_B T_{\text{CMB}}}} - 1}{e^{\frac{h\nu_{\text{ref}}}{k_B T_{\text{CMB}}}} - 1} \right)^2 \left( \frac{e^{\frac{h\nu_{\text{ref}}}{k_B T_{\text{MBB}}}} - 1}{e^{\frac{h\nu}{k_B T_{\text{MBB}}}} - 1} \right), \quad (9)$$

这里  $T_{\text{MBB}} \approx 20 \text{ K}$  是 MBB 谱的等效温度； $T_{\text{CMB}} \approx 2.726 \text{ K}$  是 CMB 背景温度； $\beta_d$  是频谱指数； $\nu_{\text{ref}}$  是可以任意指定的参照频率。因为尘埃热辐射是高频段的主要前景，我们经常选取一个比较高的参照频率，例如  $\nu_{\text{ref,dust}} \sim 300 \text{ GHz}$ 。

银河系电子同步辐射近似服从幂率谱，以 CMB 背景温度为参照黑体辐射谱来描述等效的温度扰动，幂率谱可以写作：

$$W_s(\nu; \nu_{\text{ref}}, \beta_s) = \left( \frac{\nu}{\nu_{\text{ref}}} \right)^{\beta_s-2} e^{\frac{h(\nu_{\text{ref}}-\nu)}{k_B T_{\text{CMB}}}} \left( \frac{e^{\frac{h\nu}{k_B T_{\text{CMB}}}} - 1}{e^{\frac{h\nu_{\text{ref}}}{k_B T_{\text{CMB}}}} - 1} \right)^2, \quad (10)$$

这里  $\beta_s$  是频谱指数； $\nu_{\text{ref}}$  是可以任意指定的参照频率（虽然为了节省篇幅使用了同一个符号，它和尘埃的参照频率不是同一个参数）。因为同步辐射在低频端占主导，我们经常选取一个低的参照频率，例如  $\nu_{\text{ref,sync}} \sim 30 \text{ GHz}$ 。

如果没有引起混淆的危险，我们有时把  $W_d(\nu; \nu_{\text{ref}}, \beta_d)$  简写为  $W_d(\nu)$ ，把  $W_s(\nu; \nu_{\text{ref}}, \beta_d)$  简写为  $W_s(\nu)$ 。如果  $\beta_d$ （或  $\beta_s$ ）是已知的，我们可以通过乘以  $\frac{W_d(\nu_2)}{W_d(\nu_1)}$ （ $\frac{W_s(\nu_2)}{W_s(\nu_1)}$ ）的因子把尘埃（同步辐射）在频率  $\nu_1$  的天图转化为在频率  $\nu_2$  的天图。（请注意这个比例因子和参照频率  $\nu_{\text{ref}}$  的选择是无关的。）最后，对一个有限宽的频段（假设归一化的频率分布函数是  $f(\nu)$ ），应该用积分  $\int W_d(\nu) f(\nu) d\nu$  替换  $W_d(\nu)$ 。

DroPS 默认把时序数据滤波处理的尘埃和同步辐射天图在每个  $\ell$  区间的带功率都当作自由参数。这样做有两个好处：第一，我们直接参数化滤波后天图的带功率，跳过了对时序数据滤波的建模（ $F_{\ell\ell'}$ ），避免了因建模不准确引起的系统误差。第二，虽然足够大天区里的前景功率谱一般是光滑的，在以测量原初 B 模为目的的深度观测小天区里，前景的带功率有可能存在类似于宇宙方差的涨落；在每个  $\ell$  区间内取独立的带功率就自然地把这种涨落考虑在内了。当然，这样放弃前景功率谱的光滑性假设的方法也有明显的缺陷，就是自由参数变得非常多，降低了对  $r$  的限制能力。不过，这在某种程度上恰恰体现了前景不确定性带来的系统误差必须被考虑在内。

对尘埃-尘埃、尘埃-同步辐射及同步辐射-同步辐射带功率进行盲参数化建模时，需要引入  $3n_{\text{fld}}n_{\text{bin}} + 2$  个独立参数。其中 +2 对应尘埃光谱指数  $\beta_d$  与同步辐射光谱指数



$\beta_s$ 。即便仅使用  $XY = BB$  关联场，前景参数数量通常仍不少于 20 个。这对 DroPS 中用于参数推断的马尔可夫链蒙特卡洛模拟的收敛性提出了挑战。然而，当观测数据包含能分别提供尘埃与同步辐射带功率很好的“初始估计值”的高频和低频通道时，我们发现 MCMC 模拟仍然能够实现良好收敛。

若可用频段数量有限或噪声水平较高（目标精度  $\sigma_r \gtrsim 0.01$ ），则建议采用平滑参数化方法描述前景功率谱。其原理在于此时前景的“宇宙方差”效应及滤波模型不确定性相较于噪声都已可忽略不计。通过在配置文件中设置 `analytic_fg` 为 `True`，可启用前景带功率的平滑模型——该模型以  $\ell$  的二次函数形式进行参数化。

### 3.4.2 前景能谱的空间变化：矩展开方法

以上讨论均基于  $\beta_d$  与  $\beta_s$  为常数的假设。然而对于覆盖较大天区范围的实验而言， $\beta_d$  与  $\beta_s$  不太可能在整片天区保持严格恒定 [18]。DroPS 通过矩展开近似 [19] 或泰勒展开方法处理前景光谱指数的空间变化特性，用户可在配置文件中通过设置 `freq_decorr_model="ME"` 或 `freq_decorr_model="Taylor"` 来启用相应功能。若需保持  $\beta_d$  与  $\beta_s$  严格恒定，请设置 `freq_decorr_model=None`。默认情况下，DroPS 采用矩展开方法，具体细节如下所述。

以频率为  $\nu$  的尘埃 B 模天图为例，如果没有谱指数的空间变化，在每个像素（用方向矢量  $\mathbf{n}$  来标记）上的 B 值可以写作：

$$B(\nu, \mathbf{n}) = W_d(\nu; \nu_{\text{ref}}, \beta_d) B(\nu_{\text{ref}}, \mathbf{n}), \quad (11)$$

这里  $B(\nu_{\text{ref}}, \mathbf{n})$  是参照频率上的 B 模天图（简称参照天图）。现在考虑谱指数  $\beta_d$  的微小空间波动  $\delta\beta_d(\mathbf{n})$ ，并且把上式展开到  $\delta\beta_d(\mathbf{n})$  的二阶：

$$B(\nu, \mathbf{n}) = \left\{ W_d(\nu; \nu_{\text{ref}}, \beta_d) + \frac{\partial W_d}{\partial \beta_d} \delta\beta_d(\mathbf{n}) + \frac{1}{2} \frac{\partial^2 W_d}{\partial \beta_d^2} [\delta\beta_d(\mathbf{n})]^2 \right\} B(\nu_{\text{ref}}, \mathbf{n}). \quad (12)$$

因为  $W_d \propto \left(\frac{\nu}{\nu_{\text{ref}}}\right)^{\beta_d}$ ，(12)式可以更加明确地写作

$$B(\nu, \mathbf{n}) = W_d(\nu; \nu_{\text{ref}}, \beta_d) \left\{ 1 + \delta\beta_d(\mathbf{n}) \ln \frac{\nu}{\nu_{\text{ref}}} + \frac{1}{2} \left[ \delta\beta_d(\mathbf{n}) \ln \frac{\nu}{\nu_{\text{ref}}} \right]^2 \right\} B(\nu_{\text{ref}}, \mathbf{n}). \quad (13)$$

我们首先考虑不存在观测掩膜的情况，也就是对全天天图的处理。假设  $\delta\beta_d(\mathbf{n}) = \sum \beta_{\ell m} Y_{\ell m}(\mathbf{n})$  以及  $B(\nu, \mathbf{n}) = \sum a_{\ell m}(\nu) Y_{\ell m}(\mathbf{n})$ ，把 (13) 进行球谐展开，系数为：

$$\begin{aligned} a_{\ell m}(\nu) &= W_d(\nu; \nu_{\text{ref}}, \beta_d) \int Y_{\ell m}^*(\mathbf{n}) \left\{ 1 + \delta\beta_d(\mathbf{n}) \ln \frac{\nu}{\nu_{\text{ref}}} + \frac{1}{2} \left[ \delta\beta_d(\mathbf{n}) \ln \frac{\nu}{\nu_{\text{ref}}} \right]^2 \right\} B(\nu_{\text{ref}}, \mathbf{n}) d^2\mathbf{n} \\ &= W_d(\nu; \nu_{\text{ref}}, \beta_d) \left[ a_{\ell m}^{\text{ref}} + \ln \frac{\nu}{\nu_{\text{ref}}} g_{\ell m}^{\text{ref}} + \left( \ln \frac{\nu}{\nu_{\text{ref}}} \right)^2 h_{\ell m}^{\text{ref}} \right] \end{aligned} \quad (14)$$

这里

$$a_{\ell m}^{\text{ref}} = a_{\ell m}(\nu_{\text{ref}}), \quad (15)$$

$$g_{\ell m}^{\text{ref}} = (-1)^m \sqrt{\frac{2\ell+1}{4\pi}} \sum_{\ell_1, m_1, \ell_2, m_2} a_{\ell_1, m_1}^{\text{ref}} \beta_{\ell_2, m_2} \sqrt{(2\ell_1+1)(2\ell_2+1)} \begin{pmatrix} \ell_1 & \ell_2 & \ell \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \ell_1 & \ell_2 & \ell \\ m_1 & m_2 & -m \end{pmatrix}, \quad (16)$$

以及

$$h_{\ell m}^{\text{ref}} = (-1)^m \frac{\sqrt{2\ell+1}}{8\pi} \sum_{\ell_1, m_1, \ell_2, m_2, \ell_3, m_3} a_{\ell_1, m_1}^{\text{ref}} \beta_{\ell_2, m_2} \beta_{\ell_3, m_3} \sqrt{(2\ell_1+1)(2\ell_2+1)(2\ell_3+1)} \\ \times \sum_{\ell_4, m_4} (-1)^{m_4} (2\ell_4+1) \begin{pmatrix} \ell_2 & \ell_3 & \ell_4 \\ m_2 & m_3 & -m_4 \end{pmatrix} \begin{pmatrix} \ell_2 & \ell_3 & \ell_4 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \ell_4 & \ell_1 & \ell \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \ell_4 & \ell_1 & \ell \\ m_4 & m_1 & -m \end{pmatrix}. \quad (17)$$

假设  $\delta\beta_d(\mathbf{n})$  是一个和参照天图无关联的随机高斯场，并且仅保留到  $\delta\beta_d$  的二阶小量，就有

$$\langle (a_{\ell m}^{\text{ref}})^* g_{\ell m}^{\text{ref}} \rangle = 0, \quad (18)$$

$$\langle (h_{\ell m}^{\text{ref}})^* g_{\ell m}^{\text{ref}} \rangle = 0, \quad (19)$$

以及

$$\langle (h_{\ell m}^{\text{ref}})^* h_{\ell m}^{\text{ref}} \rangle = 0. \quad (20)$$

这里  $\langle \cdot \rangle$  代表期待值。

如果进一步把参照天图  $B(\nu_{\text{ref}}, \mathbf{n})$  也近似当作随机高斯场，我们可以用参照天图的功率谱  $C_\ell^{\text{ref}} \equiv \langle (a_{\ell m}^{\text{ref}})^* a_{\ell m}^{\text{ref}} \rangle$  和  $\delta\beta_d$  的功率谱  $S_\ell \equiv \langle (\beta_{\ell m})^* \beta_{\ell m} \rangle$  来表达下述非零的期待值：

$$\langle (g_{\ell m}^{\text{ref}})^* g_{\ell m}^{\text{ref}} \rangle = \frac{1}{4\pi} \sum_{\ell_1, \ell_2} C_{\ell_1}^{\text{ref}} S_{\ell_2} (2\ell_1+1)(2\ell_2+1) \begin{pmatrix} \ell_1 & \ell_2 & \ell \\ 0 & 0 & 0 \end{pmatrix}^2, \quad (21)$$

和

$$\langle (a_{\ell m}^{\text{ref}})^* h_{\ell m}^{\text{ref}} \rangle = \frac{1}{8\pi} C_\ell^{\text{ref}} \sum_{\ell_2} (2\ell_2+1) S_{\ell_2}. \quad (22)$$

在推导(21)式的过程中，我们使用了 3-j 符号的正交定理：

$$\sum_{m_1, m_2} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} \ell_1 & \ell_2 & \ell'_3 \\ m_1 & m_2 & m'_3 \end{pmatrix} = \begin{cases} \frac{1}{2\ell_3+1}, & \text{if } |\ell_1 - \ell_2| \leq \ell_3 \leq \ell_1 + \ell_2, \ell_3 = \ell'_3, m_3 = m'_3 \\ 0, & \text{otherwise} \end{cases}. \quad (23)$$

在推导(22)式的过程中，我们使用了 3-j 符号的下述性质：

$$\sum_m (-1)^{\ell-m} \begin{pmatrix} \ell & \ell & n \\ m & -m & 0 \end{pmatrix} = \begin{cases} \sqrt{2\ell+1}, & \text{if } n = 0 \\ 0, & \text{else} \end{cases}. \quad (24)$$

根据上面的结果，两张频率天图  $B(\nu, \mathbf{n})$  和  $B(\nu', \mathbf{n})$  之间的功率谱则是

$$\begin{aligned}
C_\ell(\nu, \nu') &= \langle a_{\ell m}(\nu) a_{\ell m}(\nu') \rangle \\
&= W_d(\nu; \nu_{\text{ref}}, \beta_d) W_d(\nu'; \nu_{\text{ref}}, \beta_d) \\
&\quad \times \left\{ C_\ell^{\text{ref}} + \langle (g_{\ell m}^{\text{ref}})^* g_{\ell m}^{\text{ref}} \rangle \ln \frac{\nu}{\nu_{\text{ref}}} \ln \frac{\nu'}{\nu_{\text{ref}}} + \langle (a_{\ell m}^{\text{ref}})^* h_{\ell m}^{\text{ref}} \rangle \left[ \left( \ln \frac{\nu}{\nu_{\text{ref}}} \right)^2 + \left( \ln \frac{\nu'}{\nu_{\text{ref}}} \right)^2 \right] \right\} \\
&= W_d(\nu; \nu_{\text{ref}}, \beta_d) W_d(\nu'; \nu_{\text{ref}}, \beta_d) \\
&\quad \times \left\{ C_\ell^{\text{ref}} \left[ 1 + \frac{1}{8\pi} \left( \ln^2 \frac{\nu}{\nu_{\text{ref}}} + \ln^2 \frac{\nu'}{\nu_{\text{ref}}} \right) \sum_{\ell_2} (2\ell_2 + 1) S_{\ell_2} \right] \right. \\
&\quad \left. + \frac{1}{4\pi} \ln \frac{\nu}{\nu_{\text{ref}}} \ln \frac{\nu'}{\nu_{\text{ref}}} \sum_{\ell_1, \ell_2} C_{\ell_1}^{\text{ref}} S_{\ell_2} (2\ell_1 + 1)(2\ell_2 + 1) \begin{pmatrix} \ell_1 & \ell_2 & \ell \\ 0 & 0 & 0 \end{pmatrix}^2 \right\}, \tag{25}
\end{aligned}$$

可以把上述结果写成一个更加精简的形式：

$$C_\ell(\nu, \nu') = W_d(\nu) W_d(\nu') \left\{ C_\ell^{\text{ref}} \left[ 1 + \mu \left( \ln^2 \frac{\nu}{\nu_{\text{ref}}} + \ln^2 \frac{\nu'}{\nu_{\text{ref}}} \right) \right] + \varepsilon_\ell \ln \frac{\nu}{\nu_{\text{ref}}} \ln \frac{\nu'}{\nu_{\text{ref}}} \right\}, \tag{26}$$

这里的

$$\mu \equiv \frac{1}{8\pi} \sum_{\ell_2} (2\ell_2 + 1) S_{\ell_2} \tag{27}$$

正比于  $\delta\beta_d$  图的方差，而

$$\varepsilon_\ell \equiv \frac{1}{4\pi} \sum_{\ell_1, \ell_2} C_{\ell_1}^{\text{ref}} S_{\ell_2} (2\ell_1 + 1)(2\ell_2 + 1) \begin{pmatrix} \ell_1 & \ell_2 & \ell \\ 0 & 0 & 0 \end{pmatrix}^2 \tag{28}$$

正比于  $\beta_d$  和参照天图的卷积。

上述参数化涉及两个未知的功率谱  $C_\ell^{\text{ref}}$  和  $S_\ell$ ，而  $\mu$  和  $\varepsilon_\ell$  可以从这两个功率谱导出。

在实际的 CMB 实验中，尤其是地基实验中，我们需要处理只覆盖部分天区（天空占比  $f_{\text{sky}}$ ）的掩膜的影响。这时，(26)式、(27)式和(28)式中的功率谱都是伪功率谱（天图和掩膜乘积的功率谱），而 DroPS 处理的是带功率，也就是对统计各向同性的全天功率谱的估计。把伪功率谱仍然写作  $\tilde{S}_\ell$  和  $\tilde{C}_\ell$ ，它们和全天功率谱之间的关系为 [20]

$$\tilde{S}_\ell = \frac{2\ell' + 1}{4\pi} \sum_{\ell', \ell''} (2\ell'' + 1) \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix}^2 M_{\ell''} S_{\ell'}, \tag{29}$$

$$\tilde{C}_\ell^{\text{BB}} = \sum_{\ell'} Q_{\ell\ell'}^- C_{\ell'}^{\text{BB}} + Q_{\ell\ell'}^+ C_{\ell'}^{\text{EE}}, \tag{30}$$

$$\tilde{C}_\ell^{\text{EE}} = \sum_{\ell'} Q_{\ell\ell'}^+ C_{\ell'}^{\text{EE}} + Q_{\ell\ell'}^- C_{\ell'}^{\text{BB}}. \tag{31}$$

这里的

$$Q_{\ell\ell'}^\pm = \frac{2\ell' + 1}{4\pi} \sum_{\ell''} (2\ell'' + 1) M_{\ell''} \begin{pmatrix} \ell & \ell' & \ell'' \\ 2 & -2 & 0 \end{pmatrix}^2 \frac{1 \pm (-1)^{\ell+\ell'+\ell''}}{2}, \tag{32}$$

这里的  $M_{\ell''}$  是掩膜的功率谱。

可以看到, 掩膜让  $\mu$  和  $\varepsilon_{\ell}$  不再简单依赖于  $S_{\ell}$  和  $C_{\ell}^{BB, \text{ref}}$ , 而同时受到  $C_{\ell}^{\text{EE}, \text{ref}}$  以及  $M_{\ell}$  的影响。考虑到这些影响, 等效的  $\mu$  不再是一个尺度无关的因子。在 DroPS 中实际使用的参数化形式为

$$D_{\ell}(\nu, \nu') = W_d(\nu)W_d(\nu') \left\{ D_{\ell}^{\text{ref}} \left[ 1 + \mu_{\ell} \left( \ln^2 \frac{\nu}{\nu_{\text{ref}}} + \ln^2 \frac{\nu'}{\nu_{\text{ref}}} \right) \right] + \varepsilon_{\ell} \ln \frac{\nu}{\nu_{\text{ref}}} \ln \frac{\nu'}{\nu_{\text{ref}}} \right\}, \quad (33)$$

这里的  $D_{\ell}^{\text{ref}}$ ,  $\mu_{\ell}$  和  $\varepsilon_{\ell}$  都以带功率的形式出现并在默认情况下都被当作自由参数。为了保持原始的物理图像, 也就是  $\mu_{\ell}$  和  $\varepsilon_{\ell}$  因子的项都是小修正, DroPS 给  $|\mu_{\ell}|$  和  $|\varepsilon_{\ell}|$  设定了一个上限 (默认值为 0.05, 用户可以通过设置文件修改)。

撰写本文时, 尚未有观测证据能够证实或否定  $\beta_d$  在小尺度上存在空间变化。公式 (27,28) 表明, 若模式混合的额外贡献可被忽略, 则  $\varepsilon_{\ell}$  和  $\mu_{\ell}$  应为正参数。然而, 若  $\beta_d$  实际上不存在空间变化, 施加此物理先验可能会导致  $r$  的测量结果产生偏差, 且我们目前无法确定模式混合效应是否可被安全忽略。因此, DroPS 默认不施加正性先验。如需启用该先验, 可在配置文件中设置 `ME_is_positive = True`。

对同步辐射的处理方式完全相同, 因为  $W_d$  对  $\beta_d$  的依赖关系与  $W_s$  对  $\beta_s$  的依赖关系具有相同形式 ( $W_d \propto (\nu/\nu_{\text{ref}})^{\beta_d}$ ,  $W_s \propto (\nu/\nu_{\text{ref}})^{\beta_s}$ )。

总而言之, 我们已将矩展开方法推广至掩膜天区, 但仍忽略了实际中可能存在的诸多复杂因素, 例如  $\delta\beta_d$ 、 $\delta\beta_s$  与参考天图之间可能存在的关联性。尽管部分复杂性或可被自由参数  $\mu_{\ell}$  和  $\varepsilon_{\ell}$  吸收, 但我们不认为公式 (33) 能够涵盖所有可能的自由度。尽管如此, 至少对于基础 (恒定能谱分布) 模型的小扰动而言, 矩展开方法仍是处理前景复杂性的有效工具。

### 3.4.3 前景频率去相关: 泰勒展开方法

对于基础 (恒定能谱分布) 的尘埃 (同步辐射) 模型, 其带功率满足关系式  $[D_{\ell}(\nu, \nu')]^2 = D_{\ell}(\nu, \nu)D_{\ell}(\nu', \nu')$ 。前景的复杂性 (例如能谱分布的空间变化) 通常会减弱不同频率天图间的相关性, 即满足:

$$[D_{\ell}(\nu, \nu')]^2 \leq D_{\ell}(\nu, \nu)D_{\ell}(\nu', \nu'), \quad (34)$$

当且仅当  $\nu = \nu'$  时取等号。

但需注意的是, 矩展开公式 (33) (及其全天区版本 (26)) 并未显式遵循这一具有物理意义的不等式。受不等式 (34) 的启发, DroPS 提供了一个唯象模型, 可通过在配置文件中设置 `fg_decorr_model = "Taylor"` 启用。该模型的核心思想是对谱指数  $\beta_d$  在  $\nu, \ell$  空间进行泰勒展开。具体而言, 尘埃的带功率被参数化为:

$$D_{\ell}(\nu, \nu') = W_d(\nu; \nu_{\text{ref}}, \beta_d + \delta\beta_{d,\ell} + \gamma_d \frac{\nu}{\nu_{\text{ref}}}) W_d(\nu'; \nu_{\text{ref}}, \beta_d + \delta\beta_{d,\ell} + \gamma_d \frac{\nu'}{\nu_{\text{ref}}}) \exp \left[ -\frac{\ln^2 \frac{\nu}{\nu'}}{Q_{\ell}} \right] D_{\ell}^{\text{ref}}, \quad (35)$$

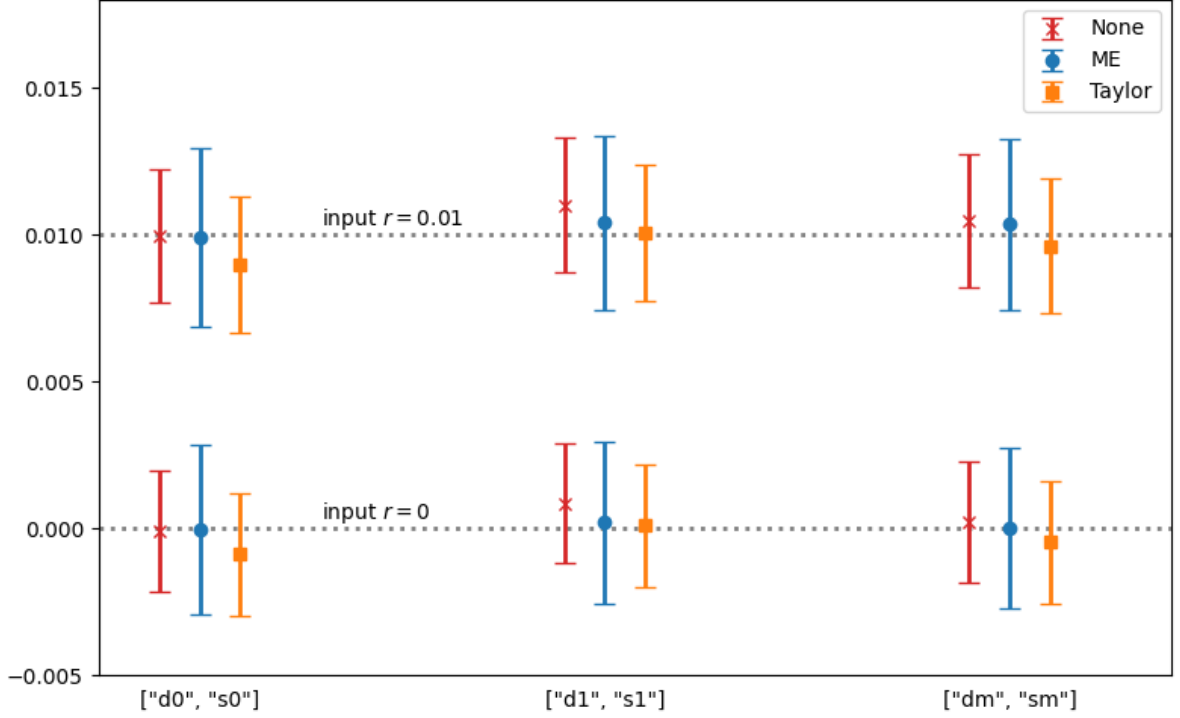


Figure 4: 基础模型 (None), 矩展开方法 (ME) 和泰勒展开方法 (Taylor) 的比较. 模拟数据基于 DroPS 的演示实验 SO。

其中  $\beta_d$  和  $\gamma_d$  分别表示平均能谱分布及其随频率的跑动;  $D_\ell^{\text{ref}}$  为参考带功率;  $\delta\beta_{d,\ell}$  描述  $\beta_d$  在  $\ell$  空间的变化;  $Q_\ell > 0$  描述频率退相关的对数间距 (当  $\ln \frac{\nu}{\nu'} \gtrsim Q_\ell$  时, 频率退相关效应显著)。

矩展开与泰勒展开方法通常会引起  $r$  值产生数量级为若干  $\times 10^{-4}$  的偏移。我们采用更精确的演示实验 SO 来突显前景建模的影响。图 4 对比了使用基础模型 (恒定能谱分布)、矩展开方法及泰勒展开方法重构的  $r$  值。每个数据点及误差棒分别代表对 250 次”观测天图”随机实现结果进行平均后得到的均值与标准差。我们的分析表明：矩展开方法在所有测试场景中均保持近乎无偏的特性，但其代价是不确定性略有增加。泰勒展开方法在前景存在频率退相关时同样表现稳健，然而若未出现此类退相关效应时，该方法则会系统地低估  $r$  值。

### 3.5 带功率的相关矩阵

虽然带功率的协方差矩阵原则上可以直接通过参照数据组估计，但这种方法面临一个实际挑战：受算力限制，通常参照数据组的样本数目在几百到几千之间，无法对包含大量自由度的协方差矩阵进行可靠估计。为解决此问题，DroPS 从物理先验的角度对协方差矩阵的计算实施两种优化方案：通过忽略不同频段或观测季节间的噪声关联来实现”噪声清理”，并通过忽略相距较远的多极矩区间之间的关联来实施”尺度分离”。

DroPS 将协方差矩阵分解为三个组成部分：噪声协方差、信号协方差以及信噪交叉

协方差。这样处理的优势在于，可以基于不同频率通道（或观测季节）的噪声互不关联这一先验知识对噪声协方差进行“清理”。DroPS 同时假设噪声的  $T$ 、 $E$ 、 $B$  分量互不关联<sup>13</sup>。因此， $D^{XY}\ell(\nu_1, \nu_2)$  与  $D^{X'Y'}\ell'(\nu'_1, \nu'_2)$  之间的噪声协方差仅当  $XY = X'Y'$  且  $(\nu_1, \nu_2) = (\nu'_1, \nu'_2)$  时才非零。这一机制显著降低了自由度数量，并提升了噪声协方差估计的准确性。

默认情况下，DroPS 假设不同多极矩区间之间不存在相关性（配置文件中设置 `ell_cross_range = 0`）。用户可选择建模相邻区间（`ell_cross_range = 1`）或次相邻区间（`ell_cross_range = 2`）的关联。然而，简单地将协方差矩阵截断至这些有限的非对角区块可能会破坏其正定性，从而引入具有负特征值的非物理“鬼模”，最终导致灾难性的过拟合现象。为平衡信息保留与数值稳定性，DroPS 对  $\ell$  非对角区块施加抑制因子来保证协方差矩阵的正定性。通过解析或者数值计算，我们计算了这些抑制因子的理论最优值，并列于表 3 中。具体来说，它们的使用方式是：

- 对 `ell_cross_range = 0`，同一  $\ell$ -区间关联不动，其余元素清零。
- 对 `ell_cross_range = 1`，同一  $\ell$ -区间关联不动，相邻  $\ell$ -区间的关联乘以  $f_1$ ，其余元素清零。
- 对 `ell_cross_range = 2`，同一  $\ell$ -区间关联不动，相邻  $\ell$ -区间的关联乘以  $f_1$ ，次相邻  $\ell$ -区间关联乘以  $f_2$ ，其余元素清零。

通过对 DroPS 的演示实验的测试，我们发现相邻  $\ell$  区间或者次相邻  $\ell$  区间的关联对  $r$  的测量影响很弱。图 5 以 AliCPT 演示实验为例展示了 `ell_cross_range = 0, 1, 2` 时  $r$  和  $\beta_d$  的测量结果。图中用虚线标记了模拟采用的  $r = 0.01$  和  $\beta_d = 1.54$ 。这个图是用 DroPS 自带的画图脚本 `utils/plot_cont.py` 生成的。

信号协方差以及信噪交叉协方差都依赖于宇宙学，特别是  $r$  的值（DroPS 默认假设其他宇宙学参数的变化范围很小，对 B 模式功率谱的影响不起主要作用）。通过  $r = 0$  and  $r = r_1$  的两组参照数据，DroPS 计算出这两种情况下的信号协方差以及信噪交叉协方差<sup>14</sup>。因为信号（带功率）和  $r$  近似有线性依赖关系，所以在  $0 < r < r_1$  的范围内，DroPS 把信号协方差按照  $r^2$  权重进行插值，把信噪交叉协方差按照  $r$  权重进行插值。对  $r > r_1$ ，DroPS 就简单使用  $r = r_1$  的协方差矩阵。DroPS 不对协方差矩阵进行外插值的原因是：对两个协方差矩阵进行外插值无法保证其正定性，容易因产生负本征值而导致灾难性的过拟合。

<sup>13</sup>在使用  $BB$  带功率的默认设置中，不需要此假设

<sup>14</sup>注意参照数据组里的前景成分是固定的，因此协方差矩阵不包含前景的随机性。DroPS 通过把参照频率的前景带功率都取为自由参数并进行边缘化的方式来处理前景的随机性，并且通过矩展开或者泰勒展开的方式来估算频率去相关带来的额外误差



Table 3: 保证协方差矩阵正定性的抑制因子 ( $f_1$  用于相邻  $\ell$  区间;  $f_2$  用于次相邻  $\ell$  区间)

	ell_cross_range = 1	ell_cross_range = 2
# $\ell$ bins	$f_1$	$f_1, f_2$
2	1	-
3	$\frac{1}{\sqrt{2}}$	1, 1
4	$\frac{\sqrt{5}-1}{2}$	$\frac{\sqrt{3}}{2}, \frac{1}{2}$
5	$\frac{1}{\sqrt{3}}$	$\sqrt{2 - \sqrt{\frac{28}{3}} \cos \frac{\pi + \arccos \sqrt{\frac{27}{28}}}{3}}, \sqrt{\frac{28}{3}} \cos \frac{\pi + \arccos \sqrt{\frac{27}{28}}}{3} - 1$
6	$\frac{2}{3} - \frac{\sqrt{28}}{3} \cos \frac{\pi + \arccos \frac{1}{\sqrt{28}}}{3}$	$\sqrt{\frac{3}{2} - \frac{4\sqrt{7}}{9} \cos \frac{\pi - \arccos \frac{17}{7\sqrt{7}}}{3}} + \frac{7}{9} \cos \frac{2\pi - 2 \arccos \frac{17}{7\sqrt{7}}}{3},$ $\frac{5}{6} - \frac{\sqrt{7}}{3} \cos \frac{\pi - \arccos \frac{17}{7\sqrt{7}}}{3}$
7	$\frac{\sqrt{2-\sqrt{2}}}{2}$	0.74222720 , 0.25777280
8	0.53208889	0.74535599 , $\frac{1}{3}$
9	0.52573111	0.73967857 , 0.29430138
10	0.52110856	0.73205081 , 0.26794919
11	0.51763809	0.72458926 , 0.24925315
12	0.51496392	0.72430917 , 0.29789426
13	0.51285843	0.72360680 , 0.27639320
14	0.51117030	0.72095982 , 0.25989153
15	0.50979558	0.71754959 , 0.24694521
16	0.50866092	0.71688142 , 0.28311858
17	0.50771331	0.71707895 , 0.26835006
18	0.50691364	0.71592096 , 0.25627141
19	0.50623256	0.71404546 , 0.24626867
20	0.50564767	0.71340854 , 0.27515189
$\infty$	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}, \frac{1}{4}$

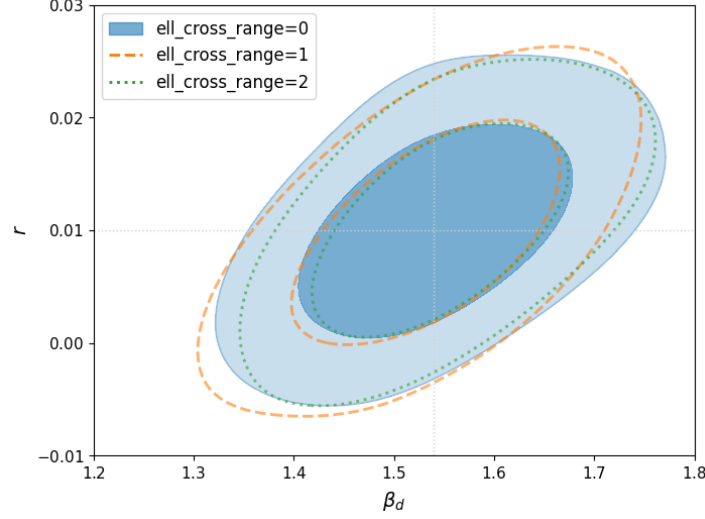


Figure 5: Impact of the correlations between adjacent  $\ell$  bins and next-to-adjacent  $\ell$  bins

### 3.6 天图成分分离的似然函数

天图成分分离涉及的噪声似然函数  $\mathcal{L} \propto e^{-\chi^2/2}$  以高斯近似的形式给出：

$$\chi^2 = v_{\text{noise}}^T N^{-1} v_{\text{noise}}, \quad (36)$$

这里的矢量  $v_{\text{noise}}$  代表所有频段上的噪声天图， $N$  代表噪声的协方差矩阵。在像素空间  $v_{\text{noise}}$  的维度为像素数目和频段数目的乘积  $n_{\text{pix}} n_\nu$ ，对类似于 AliCPT 或者 SO-SAT 等覆盖天区  $\gtrsim 10\%$  的实验而言，通常要超过  $10^4$ 。这使得协方差矩阵  $N$  难以用数值方法估算，并且它的求逆也需要消耗大量计算资源。

为了避开这一困难，BICEP/Keck 实验的 B 模天图分离丢弃了所有  $N$  的非对角元素，也就忽略了所有像素之间的噪声关联 [12]。显然地，忽略噪声的大角尺度关联会直接影响从分离出来的天图中评估出来的原初引力波的大小，因此这种做法存在一定的缺陷。DroPS 则采用不同的逻辑，以尽可能保证分离出来的 CMB 天图中的 B 模在  $\ell \lesssim 200$  范围内的准确度为目标。因此 DroPS 在谐波空间进行天图的分离。为了便于在个人电脑上测试，DroPS 默认采用  $\ell_{\text{max}} = 160$  的截断来保证有限计算资源下的可执行性。

用下标  $j$  ( $j = 1, 2, \dots, n_{\text{pix}}$ ) 和  $k$  ( $k = 1, 2, \dots, n_\nu$ ) 分别来标记像素和频段。CMB 天图在所有频段的极化天图是一样的：

$$(Q \pm iU)_{\text{CMB},j,k} = - \sum_{\ell=0}^{\ell_{\text{max}}} \sum_{m=-\ell}^{\ell} (c_{\ell m}^E \pm i c_{\ell m}^B) {}_{\pm 2}Y_{\ell m}(\mathbf{n}_j), \quad (37)$$

这里  $\mathbf{n}_j$  是  $j$  号像素对应的方向矢量。展开基  ${}_{\pm 2}Y_{\ell m}$  是自旋加权球谐函数。注意这里  $c_{\ell m}^E$  和  $c_{\ell m}^B$  是全天 CMB 图的球谐展开系数。当对掩膜覆盖天区的天图进行成分分离时，这些球谐系数之间存在简并。理论上讲，在掩膜覆盖区选取正交完备基才是计算效率最高



的做法。但是这样的代码会变得异常复杂，需要对每一个掩膜进行求解，也缺乏现成的特殊函数库支持，因此 DroPS 放弃了这种理想上可行但实际操作困难的方案。

由于频段数有限，BICEP/Keck 在做天图成分分离时仅考虑了尘埃热辐射前景。为了更好地应用到下一代 CMB 实验中，DroPS 采用了更为精确的双成分模型（尘埃热辐射和同步辐射）。忽略频率去相关，尘埃热辐射极化天图可以写作：

$$(Q \pm iU)_{\text{dust},j,k} = - \int W_d(\nu) f_k(\nu) d\nu \sum_{\ell=0}^{\ell_{\max}} \sum_{m=-\ell}^{\ell} (d_{\ell m}^E \pm i d_{\ell m}^B) {}_{\pm 2}Y_{\ell m}(\mathbf{n}_j), \quad (38)$$

这里的  $W_d$  由 (9) 式给出， $f_k(\nu)$  是第  $k$  个频段的（归一化的）频率分布函数。类似地，同步辐射天图可以写作：

$$(Q \pm iU)_{\text{sync},j,k} = - \int W_s(\nu) f_k(\nu) d\nu \sum_{\ell=0}^{\ell_{\max}} \sum_{m=-\ell}^{\ell} (s_{\ell m}^E \pm i s_{\ell m}^B) {}_{\pm 2}Y_{\ell m}(\mathbf{n}_j), \quad (39)$$

这里  $W_s$  由 (10) 式给出。

天图模型  $(Q \pm iU)_{\text{sky}}$  是公式 (37, 38, 39) 的叠加。使天图模型通过时序数据滤波及天图制作流程后，我们得到滤波后的天图。将滤波后的天图从观测到的（同样经过滤波后的）天图中扣除，便得到像素空间的滤波图。为计算似然函数，我们将噪声图分解至谐波空间：

$$M_j(Q \pm iU)_{\text{noise,filtered},j,k} = - \sum_{\ell=0}^{\ell_{\max}} \sum_{m=-\ell}^{\ell} (\tilde{n}_{\ell m}^E \pm i \tilde{n}_{\ell m}^B) {}_{\pm 2}Y_{\ell m}(\mathbf{n}_j). \quad (40)$$

此处  $M_j$  为平滑掩膜的像素值，其边缘采用 C2 型核函数并以 2 度为尺度进行平滑处理 [15]。引入此类平滑掩膜的目的是抑制由锐利边缘导致的非物理模式，并减少  $E$ - $B$  泄漏。

另一方面，我们可以从参照数据组中的噪声图计算  $\tilde{n}_{\ell m}^E$  和  $\tilde{n}_{\ell m}^B$ ，并计算其对角协方差  $\tilde{N}_{\ell m}^E \equiv \langle |\tilde{n}_{\ell m}^E|^2 \rangle$  与  $\tilde{N}_{\ell m}^B \equiv \langle |\tilde{n}_{\ell m}^B|^2 \rangle$ 。对于大多数掩膜形状而言， $\tilde{N}_{\ell m}^E \equiv \langle |\tilde{n}_{\ell m}^E|^2 \rangle$  和  $\tilde{N}_{\ell m}^B \equiv \langle |\tilde{n}_{\ell m}^B|^2 \rangle$  对  $m$  的依赖性主要反映的是有限采样导致的统计涨落而非物理内容。因此，DroPS 采用对  $m$  平均后的伪功率谱  $\tilde{N}_{\ell}^E$  和  $\tilde{N}_{\ell}^B$  来更稳健地构建谐波空间中的似然函数。该似然函数  $\propto e^{-\chi^2/2}$  由下式给出：

$$\chi^2 = f_{\text{sky}} \sum_{\ell=0}^{\ell_{\max}} \sum_{m=-\ell}^{\ell} \frac{|\tilde{n}_{\ell m}|^2}{N_{\ell}}. \quad (41)$$

由于  $\tilde{n}_{\ell m}$  对  $c_{\ell m}$ 、 $d_{\ell m}$  和  $s_{\ell m}$  的依赖关系均为线性，因此针对给定的掩膜和天图制作管线，可以预先计算梯度模板。

在真实实验的数据分析中，计算梯度模板的任务涉及从时序数据到天图的制作（而非采用 DroPS 提供的天图制作模拟器），对计算资源的需求会非常高。不过，由于该任务对每组  $\ell, m$  模式而言是相互独立的，因此具有高度可并行性，能够高效地在计算集群上执行。

## References

- [1] Nabila Aghanim, Yashar Akrami, Mark Ashdown, et al. Planck 2018 results-vi. cosmological parameters. *Astronomy & Astrophysics*, 641:A6, September 2020.
- [2] M. Tristram, A. J. Banday, K. M. Górski, R. Keskitalo, C. R. Lawrence, K. J. Andersen, R. B. Barreiro, J. Borrill, L. P. L. Colombo, H. K. Eriksen, R. Fernandez-Cobos, T. S. Kisner, E. Martínez-González, B. Partridge, D. Scott, T. L. Svalheim, and I. K. Wehus. Improved limits on the tensor-to-scalar ratio using BICEP and Planck data. *Physical Review D*, 105(8):083524, April 2022.
- [3] John Ellis and David Wands. Inflation (2023). *arXiv e-prints*, page arXiv:2312.13238, December 2023.
- [4] P. A. R. Ade, Z. Ahmed, M. Amiri, et al. BICEP/Keck XV: The BICEP3 Cosmic Microwave Background Polarimeter and the First Three-year Data Set. *Astrophysical Journal*, 927(1):77, March 2022.
- [5] Hong Li, Si-Yu Li, Yang Liu, Yong-Ping Li, and Xinmin Zhang. Tibet’s window on primordial gravitational waves. *Nature Astronomy*, 2:104–106, February 2018.
- [6] Junzhou Zhang, Shamik Ghosh, Jiazheng Dou, Yang Liu, Siyu Li, Jiming Chen, Jiaxin Wang, Zhaoxuan Zhang, Jacques Delabrouille, Mathieu Remazeilles, Chang Feng, Bin Hu, Hao Liu, Larissa Santos, Pengjie Zhang, Wen Zhao, Le Zhang, Zhi-Qi Huang, Hong Li, and Xinmin Zhang. Forecast of Foreground Cleaning Strategies for AliCPT-1. *Astrophysical Journal Supplement Series*, 274(2):26, October 2024.
- [7] Yang Liu, Lei Ming, Marco Drewes, and Hong Li. AliCPT Sensitivity to Cosmic Reheating. *arXiv e-prints*, page arXiv:2503.21207, March 2025.
- [8] Kevin Wolz, Susanna Azzoni, Carlos Hervías-Caimapo, Josquin Errard, Nicoletta Krachmalnicoff, David Alonso, Carlo Baccigalupi, Antón Baleato Lizancos, Michael L. Brown, Erminia Calabrese, Jens Chluba, Jo Dunkley, Giulio Fabbian, Nicholas Galitzki, Baptiste Jost, Magdy Morshed, and Federico Nati. The Simons Observatory: Pipeline comparison and validation for large-scale B-modes. *Astronomy & Astrophysics*, 686:A16, June 2024.
- [9] Emilie Hertig, Kevin Wolz, Toshiya Namikawa, et al. The Simons Observatory: Combining cross-spectral foreground cleaning with multitracer B-mode delensing for improved constraints on inflation. *Physical Review D*, 110(4):043532, August 2024.

- [10] Kevork Abazajian, Graeme E. Addison, Peter Adshead, et al. CMB-S4: Forecasting Constraints on Primordial Gravitational Waves. volume 926, page 54, February 2022.
- [11] Sebastian Belkner, Julien Carron, Louis Legrand, Caterina Umiltà, Clem Pryke, Colin Bischoff, and CMB-S4 Collaboration. CMB-S4: Iterative Internal Delensing and r Constraints. *Astrophysical Journal*, 964(2):148, April 2024.
- [12] BICEP/Keck Collaboration, P. A. R. Ade, Z. Ahmed, M. Amiri, D. Barkats, R. Basu Thakur, et al. BICEP/Keck XX: Component-separated maps of polarized CMB and thermal dust emission using Planck and BICEP/Keck Observations through the 2018 Observing Season. *arXiv e-prints*, page arXiv:2509.21648, September 2025.
- [13] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann. HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere. *Astrophysical Journal*, 622(2):759–771, April 2005.
- [14] David Alonso, Javier Sanchez, Anže Slosar, and LSST Dark Energy Science Collaboration. A unified pseudo- $C_\ell$  framework. *MNRAS*, 484(3):4127–4151, April 2019.
- [15] J. Grain, M. Tristram, and R. Stompor. Polarized CMB power spectrum estimation using the pure pseudo-cross-spectrum approach. *Physical Review D*, 79(12):123515, June 2009.
- [16] Peter Ade, James Aguirre, Zeeshan Ahmed, et al. The Simons Observatory: science goals and forecasts. *JCAP*, 2019(2):056, February 2019.
- [17] Antony Lewis, Anthony Challinor, and Anthony Lasenby. Efficient computation of CMB anisotropies in closed FRW models. *Astrophysical Journal*, 538:473–476, 2000.
- [18] P. A. R. Ade, M. Amiri, S. J. Benton, et al. Analysis of Polarized Dust Emission Using Data from the First Flight of SPIDER. *Astrophysical Journal*, 978(2):130, January 2025.
- [19] S. Azzoni, M. H. Abitbol, D. Alonso, A. Gough, N. Katayama, and T. Matsumura. A minimal power-spectrum-based moment expansion for CMB B-mode searches. *JCAP*, 05:047, 2021.
- [20] Eric Hivon, Krzysztof M. Górski, C. Barth Netterfield, Brendan P. Crill, Simon Prunet, and Frode Hansen. MASTER of the Cosmic Microwave Background Anisotropy Power Spectrum: A Fast Method for Statistical Analysis of Large and Complex Cosmic Microwave Background Data Sets. *Astrophysical Journal*, 567(1):2–17, March 2002.