UNICAMP

Marcelo Zimbres Silva

Procurando por assinaturas do campo magnético cósmico com wavelets

Campinas

2014

UNICAMP

Universidade Estadual de Campinas
Instituto de Física Gleb Wataghin

Marcelo Zimbres Silva

Procurando por assinaturas do campo magnético cósmico com wavelets

Tese apresentada ao Instituto de Física Gleb Wataghin da Universidade Estadual de Campinas como parte dos requisitos exigidos para obtencão do título de Doutor em Ciências.

Orientador: Ernesto Kemp

Este exemplar corresponde à versão final de tese defendida pelo aluno Marcelo Zimbres Silva e orientada pelo Professor Ernesto Kemp.

Campinas
2014

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Física Gleb Wataghin
Valkíria Succi Vicente - CRB 8/5398

Si38p

Silva, Marcelo Zimbres, 1980-
Procurando por assinaturas do campo magnético cósmico com wavelets / Marcelo Zimbres Silva. – Campinas, SP : [s.n.], 2014.

Orientador: Ernesto Kemp.
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Física Gleb Wataghin.

1. Wavelets esféricos. 2. Multipletos. 3. Campos magnéticos cósmicos. I. Kemp, Ernesto,1965-. II. Universidade Estadual de Campinas. Instituto de Física Gleb Wataghin. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Looking for signatures of the cosmic magnetic field using wavelets
Palavras-chave em inglês:
Spherical wavelets
Multiplets
Cosmic magnetic fields
Área de concentração: Física
Titulação: Doutor em Ciências
Banca examinadora:
Ernesto Kemp [Orientador]
Ivone Freire da Mota e Albuquerque
Edivaldo Moura Santos
Carola Dobrigkeit Chinellato
Orlando Luis Goulart Peres
Data de defesa: 29-08-2014
Programa de Pós-Graduação: Física

MEMBROS DA COMISSÃO JULGADORA DA TESE DE DOUTORADO DE **MARCELO ZIMBRES SILVA – RA 015308** APRESENTADA E APROVADA AO INSTITUTO DE FÍSICA "GLEB WATAGHIN", DA UNIVERSIDADE ESTADUAL DE CAMPINAS, EM 29 / 08 / 2014.

COMISSÃO JULGADORA:

Prof. Dr. Ernesto Kemp - Orientador do Candidato
DRCC/IFGW/UNICAMP

Profa. Dra. Ivone Freire da Mota e Albuquerque – IF/USP

Prof. Dr. Edivaldo Moura Santos – IF/USP

Profa. Dra. Carola Dobrigkeit Chinellato – DRCC/IFGW/UNICAMP

Prof. Dr. Orlando Luis Goulart Peres – DRCC/IFGW/UNICAMP

Abstract

Due to the action of the intervening cosmic magnetic fields, the trajectories of ultra-high energy cosmic rays (UHECRs) can be deflected in such a way as to create clustered energy-ordered filamentary structures in the arrival directions of these particles, the so-called multiplets. In this work we propose a new method based on the spherical wavelet transform to identify multiplets in sky maps containing arrival directions of UHECRs. The method is illustrated in simulations with a multiplet embedded in isotropic backgrounds with different numbers of events, and on data from the Pierre Auger Observatory. The efficiency of the algorithm is assessed through the calculation of Type I and II errors.

Resumo

Devido à ação do campo magnético cósmico, trajetórias de raios cósmicos provenientes de uma mesma fonte podem ser defletidas e dar origem a estruturas filamentares cujos eventos são ordenados de acordo com suas energias, os multipletos. Nesse trabalho, propomos um novo método para identificação de multipletos, baseado no uso de uma classe de funções definidas sobre a esfera, chamadas de wavelets esféricos.

Para testar o método aplicamos a análise em dados simulados. Primeiramente usamos um fundo isotrópico, onde um multipleto pode ocorrer apenas ao acaso. Posteriormente fazemos a análise colocando um multipleto em uma posição aleatória no mesmo fundo isotrópico. Com isso calculamos erros de tipo I e II. O método também é aplicado em dados obtidos pelo Observatório Pierre Auger para eventos com energia $E > 15 \times 10^{18}$ eV.

Sumário

Agradecimentos

Agradeço ao meu orientador Ernesto Kemp por me propor um projeto tão interessante e adequado ao meu perfil de pesquisa. Sem sua confiança e incentivo esse trabalho certamente não teria sido possível.

Agradeço a Rafael Alves Batista que acompanhou de perto praticamente todas as fases desse trabalho. As intensas discussões que tivemos me ajudaram muito com o entendimento do problema.

Agradeço a todos os membros da Colaboração Pierre Auger e em particular aos membros do grupo brasileiro e da Universidade de Wuppertal. Dentre eles destaco Vitor de Souza, por disponibilizar o *cluster* da USP São Carlos e Karl-Heinz Kampert pelo excelente ambiente de trabalho que me proporcionou em seu grupo.

Agradeço a minha mãe Eliana Z. Silva e a irmã Fernanda Z. Silva por todo apoio dado ao longo dos anos de doutorado.

Agradeço o apoio financeiro da Capes.

Lista de Figuras

Capítulo 1

Raios cósmicos

Those who learn and do not teach
are thieves.
DESCONHECIDO

Le défault unique de tous les ouvrages
c'est d'être trop longs.
VAUVENARGUES

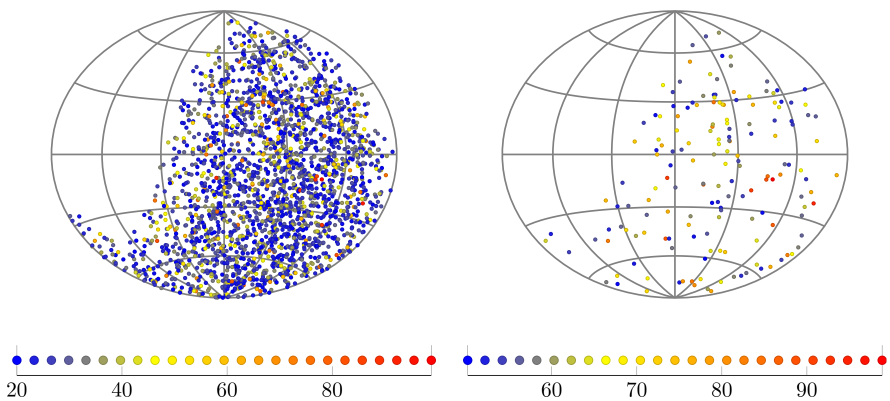
O CONCEITO de raios cósmicos é abstrato para os que se deparam com ele pela primeira vez, mas não mais para os físicos, que os estudam desde sua descoberta pelo cientista Victor Hess no começo do século XX [1]. Os avanços em técnicas de detecção feitos em mais de cem anos de pesquisa ativa, possibilitaram experimentos como Agasa [2, 3], Hires [6] e o Observatório Pierre Auger [7]. Esses experimentos contribuíram de forma decisiva para o aumento da amostragem. Dessa forma, previsões teóricas puderam ser testadas e consequentemente os modelos físicos sofreram melhorias.

Apesar de tudo isso, ainda não somos capazes de responder algumas das perguntas mais fundamentais, principalmente quando essas se referem a raios cósmicos com energias ultra altas, ou, nos termos usados pelos físicos, *Ultra High Energy Cosmic Rays* (UHECRs) [5]. Essa nomenclatura se justifica quando olhamos para o espectro de energia dos raios cósmicos. A energias de aproximadamente 100 MeV, já é possível detectar raios cósmicos, que nesse caso se originam no Sol. No outro extremo do espectro estão energias que podem superar 10^{18} eV = 1 EeV. Enquanto que na região de baixas energias o fluxo de raios cósmicos é alto o suficiente para permitir estudos com alta acurácia, na região de energias ultra altas, i.e. da ordem de EeV, os fluxos medidos na Terra são extremamente baixos, por exemplo, nessa faixa de energia, apenas uma partícula é esperada por quilômetro quadrado por ano.

O baixo fluxo de UHECRs dificulta muito sua detecção, se tivermos a esperança de detectar uma quantidade modesta de partículas diariamente, nossos detectores terão que monitorar áreas muito grandes. Por exemplo, para termos uma média de uma partícula detectada por dia, precisamos de uma área de de 365 km^2.

O baixo fluxo não é a única dificuldade encontrada pelos cientistas que projetam observatórios de UHECRs. Os raios cósmicos interagem com as primeiras camadas da atmosfera gerando partículas carregadas, como píons por exemplo. Sua identificação na superfície da Terra é possível apenas por detecção de efeitos indiretos, como ionização e a formação de chuveiros atmosféricos [11]. Para energias da ordem de GeV, os chuveiros não são intensos o suficiente para serem medidos na superfície da Terra e por isso, os detectores devem ser instalados em regiões altas, seja por meio de balões atmosféricos ou satélites. Apenas para partículas com energia da ordem de 10^{15} eV, as partículas secundárias que compõem o chuveiro atmosférico são numerosas o suficiente para serem detectadas na superfície. O chuveiro produzido nessa faixa de energia pode cobrir alguns quilômetros quadrados.

Nesse trabalho focaremos no estudo de um fenômeno físico, associado aos raios cósmicos, que acreditamos estar presente nos dados, ainda que de forma fraca, os multipletos [12]. Multipletos são alinhamentos de eventos, provenientes de uma mesma fonte, que devido à ação do campo magnético



(a) $20 < E < 100$ (b) $50 < E < 100$

Figura 1.1: Mapa de eventos dos dados do Observatório Pierre Auger para faixas de energia diferentes. Na figura da direita podemos ver que o número de partículas cai drasticamente conforme a energia cresce. As únidade de energia usada na figura é EeV.

que permeia o espaço, sofrem deflexão e portanto se alinham. Essas deflexões só são possíveis para raios cósmicos que tenham carga elétrica diferente de zero, uma vez que partículas neutras como os nêutrons e fótons não são defletidas. Dessa forma, a composição dos raios cósmicos tem consequências diretas na formação de multipletos.

As informações mais importantes do ponto de vista físico, que os experimentos nos fornecem sobre um raio cósmico são: sua direção de chegada e a energia estimada da partícula. Com esses dados e adicionalmente usando o instante em que elas nos alcançam, devemos ser capazes de descobrir de onde ela veio. O problema em um primeiro momento parece simples, pois bastaria olhar na mesma direção para ver se existe algum objeto astronômico que pudesse emitir as partículas em nossa direção. As partículas que detectamos são em sua maior parte carregadas [14], o que significa que suas trajetórias não seriam linhas retas, caso elas sofram a ação de algum campo magnético. Concluímos assim que precisamos de outra forma de inferir a posição da fonte. Além disso é possível que o objeto astronômico que esteja emitindo as partículas não esteja catalogado. Nesse sentido detectar um alinhamento de eventos pode fornecer muita informação sobre a possível localização da fonte.

Como estamos lidando com campos magnéticos muito fracos e com energias muito altas, apesar da grande distância das fontes, esperamos deflexões angulares pequenas, onde o ângulo de deflexão δ sofrido pela partícula, é proporcional ao inverso da energia E da partícula i.e. $\delta \propto 1/E$ [15].

Portanto, se formos capazes de identificar um conjunto de eventos como pertencentes a um multipleto será possível em princípio reconstruir a posição da fonte que os emitiu. A reconstrução da direção da fonte, por sua vez, é dependente do modelo de campo magnético adotado. Uma vez que tenhamos as coordenadas da fonte, poderemos comparar nossos mapas de raios cósmicos com catálogos.

O problema se reduz portanto a como identificar os multipletos, tendo em vista que estarão imersos em um fundo de eventos, que poderá ser considerado ruído. No momento em que escrevo, o número de eventos coletados pelo Observatório Pierre Auger, cujas energias são maiores que 15 EeV, é de aproximadamente 4200 eventos. Como esperamos algo em torno de 10 - 20 eventos para um multipleto [12], temos que a fração entre os eventos no multipleto e os eventos no fundo é algo da ordem de 10^{-2}.

A figura 1.2 ilustra o alinhamento de eventos de um multipleto no plano tangente a uma esfera e a figura 1.3 mostra o zoom desse plano tangente, com alguns detalhes que serão importantes ao longo desse trabalho.

1.1 Propagação

Devido ao fato de ainda não conhecermos a composição dos raios cósmicos a altas energias, temos dificuldades em falar sobre sua propagação, principalmente no que diz respeito a sua interação com campos presentes ao longo

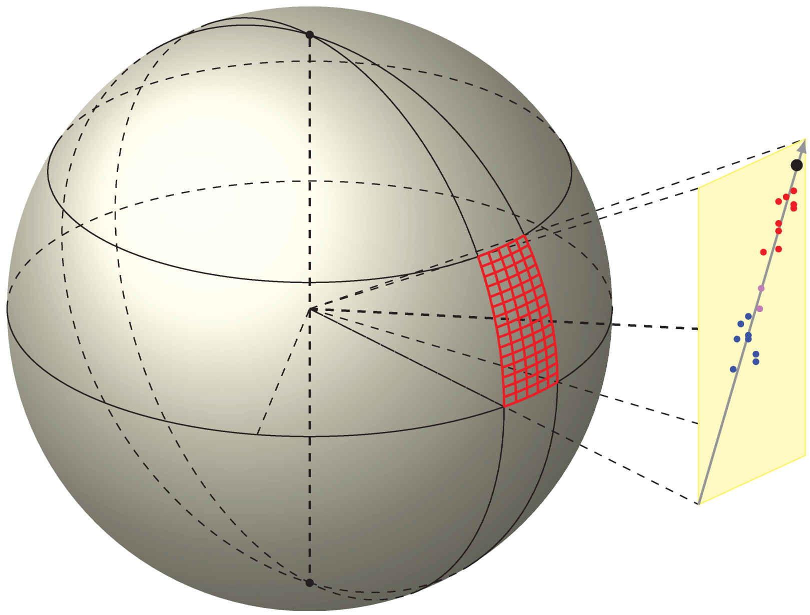


Figura 1.2: Ilustração de um multipleto na esfera celeste.

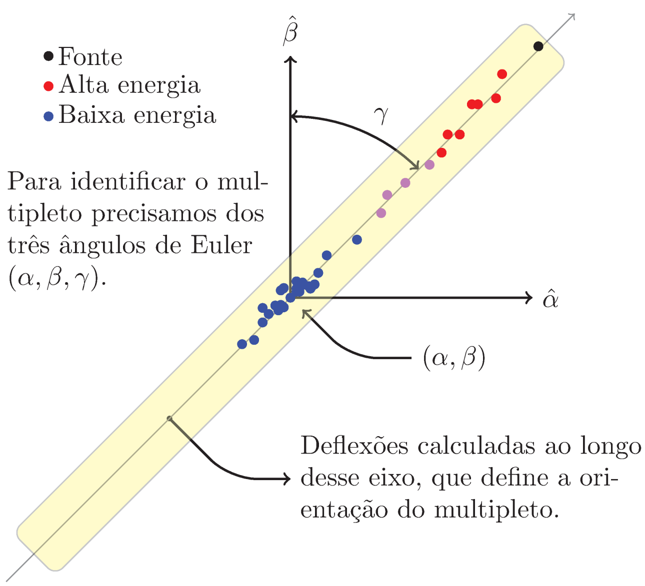


Figura 1.3: Ângulos de Euler que definem a posição e orientação do multipleto.

de sua trajetória.

É conveniente dividir o problema em duas partes, o domínio clássico e o quântico. No domínio clássico, temos apenas dois campos a considerar, o campo gravitacional, descrito pelas equações da relatividade geral e o campo eletromagnético, descrito pelas equações de Maxwell. A interação de um raio cósmico com o campo gravitacional é importante provavelmente apenas nas regiões onde exista um grande acúmulo de massa, como buracos negros, ao redor de estrelas de nêutrons e no centro da galáxia.

O efeito de lentes gravitacionais pode ser importante na localização das fontes, principalmente quando as partículas forem fótons, que por não terem carga elétrica, não podem ser defletidos por um campo magnético e mesmo assim podem dar origem a multipletos. No entanto não consideraremos esse tipo de interação nesse trabalho.

A interação de uma partícula carregada com o campo elétrico e magnético, dada pela força de Lorentz

$$\frac{d\vec{p}}{dt} = e(\vec{E} + \vec{v} \times \vec{B}), \tag{1.1}$$

é de longe a interação clássica mais importante dos raios cósmicos. Na fórmula acima e denota a carga elétrica da partícula, $\vec{p}$ seu momento linear, $\vec{v}$ sua velocidade e $\vec{E}$ e $\vec{B}$ os campos elétricos e magnéticos, respectivamente. Como o campo magnético não pode mudar a energia de uma partícula pontual carregada, é provavelmente o campo elétrico o grande responsável pela aceleração das partículas. Já nos meios inter e extragalácticos, a densidade de carga é próxima de zero, e portanto a interação mais importante é devida ao campo magnético, que deflete a trajetória da partícula.

Já do ponto de vista quântico temos várias interações possíveis, sendo a mais importante delas a interação com os fótons da Radiação Cosmológica de Fundo, que dá origem ao efeito GZK.

1.2 Espectro e composição

Medir a composição da partícula primária é considerado um dos maiores desafios em física de raios cósmicos. Como a partícula primária (nome dado ao raio cósmico que origina o chuveiro atmosférico) não é observada diretamente, a composição deve ser inferida a partir dos produtos da colisão da partícula primária com a atmosfera.

A composição varia muito com faixa de energia observada. A energias da ordem de 100 MeV os raios cósmicos têm origem unicamente solar, já que o vento solar é capaz de repelir partículas com energias dessa ordem, originárias de outras fontes. A energias mais altas, o espectro é muito bem descrito por leis de potência, ou seja, o espectro de energia é proporcional a $E^{-\gamma}$, com valores de γ que dependem da faixa de energia. A energias da ordem de $E \approx 10^{-15}$ EeV temos o joelho, chamado *knee* na literatura internacional. Nessa região de energia, o fluxo de partículas por área, tempo, ângulo sólido e energia, sofre uma mudança de $\gamma = 2.7$ para 3.0. Acima do joelho o espectro tem uma nova mudança no índice para $\gamma = 3.3$, a uma energia de $E = 4 \times 10^{17}$ EeV, que é chamado de segundo joelho. Evidências experimentais indicam que existe uma mudança na composição, passando prótons para núcleos pesados nessas energias. A última mudança que podemos detectar no índice γ, ocorre a uma energia de $E \approx 5 \times 10^{18}$ EeV, onde o índice muda para $\gamma = 2.8$.

Acredita-se que nessa faixa do espectro a maior contribuição para o fluxo venha de raios cósmicos extragalácticos.

1.3 Experimentos em raios cósmicos

Devido ao espectro tão amplo dos raios cósmicos, as técnicas de detecção variam muito. No caso de UHECRs, como mencionado na introdução, a partícula primária dificilmente poderá ser detectada, já que o fluxo de raios cósmicos nessa faixa de energia é baixíssimo e as partículas primárias podem ser detectadas apenas indiretamente.

Dois do experimentos mais importantes em raios cósmicos são:

HiRes. Hires é a sigla para *High Resolution Fly's Eye.* O experimento operou no deserto do estado do Utah nos EUA de maio de 1997 a abril de 2006 e consistia de dois conjuntos de telescópios de luz fluorescente chamados de Hires I e Hires II. Hires I consistia de um anel de 22 telescópios e podia detectar feixes de luz fluorescente, com elevação entre 3 e 17 graus. Hires II consistia de dois conjuntos de telescópios e podia observar elevações mais altas, entre 3 - 31 graus. A luz fluorescente é emitida por átomos de nitrogênio excitados pela ionização das partículas no chuveiro atmosférico, formado pela colisão da partícula primária com a atmosfera.

É atribuída ao Hires a primeira observação do efeito GZK [4].

Pierre Auger. O Observatório Pierre Auger é atualmente o maior experimento relacionado à física de raios cósmicos. Planejado inicialmente para operar nos dois hemisférios, apenas a parte sul, localizada em Pampa Amarilla na província de Mendoza, Argentina, está pronta e operante. Seu foco está na detecção de partículas com energias ultra altas i.e. da ordem de EeV e acima. O experimento usa uma técnica mista de detecção que usa tanto detectores de luz Cherenkov, quanto detectores de fluorescência. Seus 1600 detectores Cherenkov estão distribuídos em uma área de 3000 km^2. Cada um dos detectores contém 12.000 litros de água super-pura e cada um possui três fotomultiplicadoras responsáveis por detectar a luz Cherenkov.

Os 27 detectores de fluorescência são operados apenas em noites claras, quando é possível fotografar as interações causadas pela partícula primária com as primeiras camadas da atmosfera.

1.4 Multipletos

No capítulo 4 veremos os desafios que a busca de multipletos apresenta, assim como um algoritmo desenvolvido nesse trabalho. No entanto, antes disso é

Figura 1.4: Um dos detectors de superfície do Observatório Pierre Auger.

importante revisar os trabalhos mais importante já desenvolvidos sobre o assunto.

Aparentemente, a primeira proposição formal de um método para identificação de multipletos foi feita por Pierre Billoir em uma publicação restrita aos membros da Colaboração Pierre Auger com título *Searching for Threadlike Multiplets*. A principal contribuição desse trabalho está na proposição de observáveis que podem ser usados para identificar um multipleto.

A definição de multipletos, proposta pelo autor, em suas próprias palavras é a seguinte

A threadlike multiplet may be defined as a set of three or more points, each one defined by a unit vector $\vec{n}_i$, in a structure which:

- *is small: that is, for any pair,* $\vec{n_i} \cdot \vec{n_j} > \cos_m in$*: here we take* $\cos_m in =$

0.98, which corresponds to an angular distance $\Delta = 0.2$ rad ($\approx$ 11 degrees);

- *has a "length" significantly larger than its "width"; ...*

Com essas definições, o trabalho segue com a definição de quatro observáveis. O primeiro deles é a área de polígonos formados pelos eventos pertencentes ao multipleto. Segundo o autor, a área formada pelos eventos pertencentes a multipletos deverá ser menor que a área formada por eventos pertencentes a um fundo isotrópico.

O segundo observável proposto são os *parâmetros de forma*, que definem o comprimento L e a largura W de um multipleto como as raízes quadradas dos autovalores do *tensor de inércia*. O tensor de inércia é calculado no plano tangente onde o multipleto está localizado.

Com essa definição, pode-se comparar as distribuições de L e W com as distribuições calculadas a partir de fundos isotrópicos, e a partir daí, obter a significância do resultado.

O terceiro observável proposto, que também foi usado nesse trabalho, é a correlação entre o ângulo de deflexão do multipleto e a energia da partícula. No caso em que os ângulos de deflexão δ das partículas causados pelo campo magnético são pequenos, podemos demonstrar que eles são proporcionais ao inverso da energia da partícula, portanto o gráfico $\delta \times 1/E$ deve ser uma linha reta. A correlação mede o quão próximo um conjunto de pontos está de uma linha reta.

O quarto e último observável proposto na *GAP Note* (nota interna da colaboração) é o número de eventos em tiras, onde cada tira pode ser pensada como um retângulo tangente à esfera em um dado ponto P. Como multipletos são acúmulos de eventos ao longo de estruturas lineares, deveria haver tiras cujo número de eventos difere consideravelmente do caso isotrópico.

Após aplicar a análise proposta aos dados do Observatório Pierre Auger, seu trabalho é concluído sem encontrar evidências da presença de multipletos nos dados.

Outro trabalho importante sobre multipletos foi publicado oficialmente pela colaboração Pierre Auger [10]. Esse trabalho não tem como foco principal a metodologia de identificação de multipletos. O texto abaixo retirado do artigo é o único trecho que descreve o algoritmo usado.

> *We hence consider for every event above 45 EeV the quadruplets that it forms with the events within a circle of 15° having energies above 25 EeV and with a correlation coefficient* $C(u, 1/E) \geq 0.8$. *The precise values of these cuts are not crucial as long as they allow one to safely include the larger multiplets of interest. For each of these candidates we then extend the search including all the events above 20 EeV with an angular distance to the highest energy one smaller than* $20°$ *and at a distance smaller than* $3W_{max}$ *from the quadruplet axis. This allows us to find the correlated multiplets satisfying the cuts in* W_{max} *and* C_{min} *in a very efficient way, as it is desirable to be able to perform a large number of simulations.*

É interessante notar que em nenhum desses trabalhos é feita uma proposição clara de como localizar a posição e a orientação de um multipleto no céu. Apenas observáveis são propostos.

Capítulo 2

Preliminares matemáticas

Science is what we understand well
enough to explain to a computer.
Art is everything else we do.
DONALD KNUTH

Este trabalho tem como principal objetivo propor um algoritmo para busca de multipletos em mapas de eventos de raios cósmicos. O principal diferencial em relação a outros trabalhos está no uso de wavelets esféricos. Wavelets esféricos são funções com domínio em S^2, ou seja, funções do tipo $f : S^2 \to \mathbb{R}$, que obedecem a alguns critérios, como, por exemplo, ter média zero e formar uma base completa. Para que o leitor se sinta confortável com os conceitos que serão expostos nos próximos capítulos, faremos uma revisão da matemática envolvida.

Normalmente lidamos com análise de dados em uma dimensão, como, por exemplo, séries temporais, e também em duas dimensões, como imagens em geral. Mas alguns experimentos como Pierre Auger [7], CoBE [8] e WMAP [9], coletam dados em todas as direções e portanto as imagens formadas não estão mais definidas no plano, mas sim na superfície de uma esfera.

A necessidade de processar dados com essa característica, não se limita à física e pode ser encontrada em diferentes áreas de estudo como, por exemplo, visão computacional e química quântica. O tratamento desses dados apresenta novos desafios e demanda um tratamento matemático mais complexo.

Felizmente, boa parte do que precisamos rever, como funções definidas no grupo SO(3) (o grupo de rotações em três dimensões), é usado extensivamente por físicos em várias áreas, como, por exemplo, no estudo do momento angular em mecânica clássica e quântica.

Gostaríamos de enfatizar que o tratamento matemático feito aqui expõe apenas o necessário para que se possa entender a análise de funções sobre a esfera e em especial, a análise de wavelets.

2.1 O grupo SO(3)

Usaremos SO(3) para denotar o grupo de rotações em três dimensões, que normalmente é representado por matrizes 3×3. Funções definidas em SO(3) têm uma expansão de Fourier análoga à de funções definidas no círculo. O estudo desse grupo é importante aqui pois, como veremos adiante, a representação de wavelets de um sinal é uma função definida em SO(3).

Ângulos de Euler. Podemos representar qualquer rotação de um corpo rígido através de três rotações individuais. A parametrização dessas rotações individuais é feita em geral pelos ângulos de Euler, que denotamos nesse trabalho por α, β e γ. Adotaremos a convenção *zyz*, ou seja, primeiro fazemos a rotação de um ângulo γ em torno do eixo $\hat{z}$, depois uma rotação de um ângulo β em torno do novo eixo $\hat{y}$ (obtido após a primeira rotação) e finalmente um rotação de um ângulo α em torno do novo eixo $\hat{z}$ (obtido após a segunda rotação). A representação das matrizes de rotação com a parametrização mencionada acima é dada por

$$u(\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{e} \quad a(\beta) = \begin{pmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{pmatrix}, \tag{2.1}$$

onde $0 \leq \alpha < 2\pi$ e $0 \leq \beta \leq \pi$. O produto das matrizes de rotação nos dá a forma final da rotação completa que denotamos por g

$$g = g(\alpha, \beta, \gamma) = u(\alpha)a(\beta)u(\gamma). \tag{2.2}$$

A parametrização das rotações em SO(3) através dos ângulos de Euler é muito conveniente pois permite que usemos a notação $f(\alpha, \beta, \gamma)$ para nos referir a uma função f que toma valores em SO(3).

2.2 Análise de Fourier em SO(3)

Com a parametrização do grupo SO(3) definida na última seção, podemos expressar o produto interno de duas funções f e $h \in$ SO(3) como

$$\langle f, h \rangle = \int_0^{2\pi} d\alpha \int_0^{\pi} \sin\beta d\beta \int_0^{2\pi} f(\alpha, \beta, \gamma)\overline{h(\alpha, \beta, \gamma)} d\gamma, \tag{2.3}$$

onde $\overline{h(\alpha, \beta, \gamma)}$ denota o conjugado complexo de $h(\alpha, \beta, \gamma)$.

Qualquer função definida em SO(3) pode ser expandida em termos das funções D de Wigner, denotadas da forma $D^l_{mn}(\alpha, \beta, \gamma)$, que formam uma base em SO(3)

$$f(\alpha, \beta, \gamma) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \sum_{n=-l}^{l} f^l_{mn} D^l_{mn}(\alpha, \beta, \gamma). \tag{2.4}$$

Essas funções satisfazem a relação de ortogonalidade

$$\langle D^l_{mn}, D^{l'}_{m'n'}\rangle = \frac{8\pi^2}{2l+1}\delta_{ll'}\delta_{mm'}\delta_{nn'}, \tag{2.5}$$

e podem ser decompostas da seguinte forma

$$D^l_{mn}(\alpha,\beta,\gamma) = e^{-im\alpha}d^l_{mn}(\beta)e^{-in\gamma}, \tag{2.6}$$

onde as funções $d^l_{mn}(\beta)$ são chamadas de *small Wigner-d functions* na literatura internacional. Falaremos mais sobre elas no apêndice A.1. Os coeficientes de Fourier da função $f(\alpha,\beta,\gamma)$ podem ser obtidos de maneira usual com produtos internos entre os elementos da base e a função f

$$\begin{aligned} f^l_{mn} &= \frac{2l+1}{8\pi^2}\langle f, D^l_{mn}\rangle, \qquad (2.7)\\ &= \frac{2l+1}{8\pi^2}\int_0^{2\pi} d\alpha \int_0^{\pi} \sin\beta d\beta \int_0^{2\pi} f(\alpha,\beta,\gamma)\overline{D^l_{mn}(\alpha,\beta,\gamma)}d\gamma. \end{aligned}$$

Uma implementação computacional das integrais acima exige uma discretização da integral. Falaremos sobre isso nas seções seguintes.

Harmônicos esféricos. Podemos também associar um operador $\Lambda(g)$ para cada $g \in \mathrm{SO}(3)$, que age em funções $f : S^2 \to \mathbb{R}$ quadraticamente integráveis, ou seja

$$\Lambda(g)f(\omega) = f(g^{-1}\omega), \tag{2.8}$$

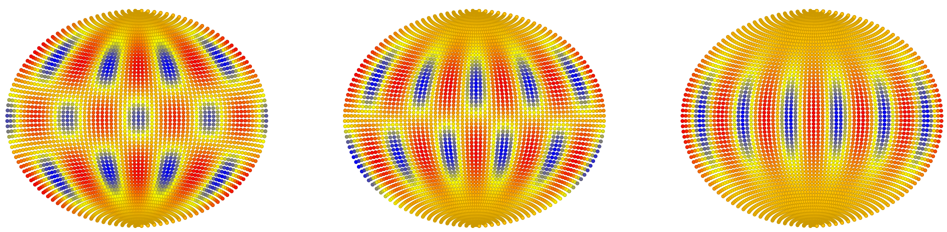
onde $\omega \in S^2$ ou usando coordenadas esféricas convencionais $\omega = (\theta,\varphi)$, onde θ é a coordenada zenital e φ azimutal. Os espaços invariantes diante dessa transformação têm como base os harmônicos esféricos Y^m_l. Portanto podemos escrever

$$\Lambda(g)Y^m_l(\omega) = \sum_{n=-l}^{l} Y^n_l(\omega)D^l_{nm}, \tag{2.9}$$

e usando explicitamente as coordenadas θ e φ

$$f(\theta,\varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_{lm} Y_m^l(\theta,\varphi), \tag{2.10}$$

onde a_{lm} são chamados de multipolos de f.



(a) (b) (c)

Figura 2.1: Harmônicos esféricos para um mesmo $l = 6$ e valores diferentes de $m = 4$, 5, 6, respectivamente.

2.3 Integração

Para calcular o produto de funções em SO(3) precisamos discretizar a integral (2.3), ou seja, transformar a integral em uma somatória finita, para isso precisamos escolher a malha onde vamos amostrar as variáveis α, β, γ. A maneira como fazemos essa discretização tem implicações muito importantes na análise de dados. Trataremos melhor esse assunto na seção 2.5. Após uma discretização da equação 2.3 a forma geral da somatória é dada por

$$\sum_{x\in X} w(x) f(x) \overline{D^l_{mn}(x)}, \tag{2.11}$$

onde X denota a malha onde amostramos a função e $w(x)$ é uma função peso. Para garantir que a somatória acima compute precisamente a transformada de Fourier de $f(x)$, precisamos impor que $f(x)$ tenha limite de banda B, onde a definição de limite de banda é dada por:

Definicão 1.

Uma função f contínua em SO(3) tem limite de banda B, se

$$f^l_{mn} = 0 \quad \forall \quad l \geq B. \tag{2.12}$$

Como estamos usando os ângulos de Euler para parametrizar as rotações, denotamos a malha na forma

$$X_B = [u(\alpha_{j_1})a(\beta_k)u(\gamma_{j_2}) \ \text{ com } \ 0 \leq j_1,\, j_2,\, k < 2B], \tag{2.13}$$

onde

$$\alpha_j = \gamma_j = \frac{2\pi j}{2B}, \quad \beta_k = \frac{\pi(2k+1)}{4B}. \tag{2.14}$$

Se impusermos que os pesos $w_B(k)$ satisfaçam a equação

$$\sum_{k=0}^{2B-1} w_B(k) P_m(\cos\beta_k) = \delta_{0m} \quad \forall \quad 0 \leq m < 2B, \tag{2.15}$$

podemos provar o seguinte teorema [19]:

Teorema 1. *Suponha que $f \in L^2(SO(3))$ é uma função com limite de banda B, então para todo $l < B$ temos*

$$f^l_{mn} = \frac{1}{(2B)^2} \sum_{j_1=0}^{2B-1} \sum_{j_2=0}^{2B-1} \sum_{k=0}^{2B-1} w_B(k) f(\alpha_{j_1}, \beta_k, \gamma_{j_2}) D^{l*}_{mn}(\alpha_{j_1}, \beta_k, \gamma_{j_2}), \tag{2.16}$$

onde os pesos têm a expressão exata

$$w_B(k) = \frac{2}{B} \sin\left(\frac{\pi(2k+1)}{4B}\right) \sum_{j=0}^{B-1} \frac{1}{2j+1} \sin\Big((2j+1)(2k+1)\frac{\pi}{4B}\Big), \tag{2.17}$$

com $0 \leq k < 2B$.

2.4 Correlação em S^2

Suponhamos que com os dados coletados de um experimento possamos formar um imagem, que, por sua vez, é definida pela função $f(\theta, \varphi)$. Chamaremos f de *sinal*. Esse sinal pode ser formado a partir das direções de incidência de raios cósmicos na Terra, por exemplo. Suponhamos também que saibamos da existência, mas não da localização, de um multipleto imerso no sinal, ou seja $f = \textit{fundo} + \textit{multipleto}$.

Se conhecemos o sinal $h(\theta, \varphi)$ que descreve o multipleto, podemos procurá-lo no sinal correlacionando todas as possíveis rotações de h com f e selecionar a que produz uma maior correlação. Portanto precisamos de uma maneira de correlacionar f e h.

Matematicamente o que precisamos fazer é correlacionar f e h da forma

$$C(\alpha, \beta, \gamma) = \int_{S^2} f(\theta, \varphi)\overline{\Lambda(\alpha, \beta, \gamma)h(\theta, \varphi)}d(\cos\theta)d\varphi, \tag{2.18}$$

e pegar os valores α, β e γ para os quais C é um máximo local. Essa abordagem no entanto demanda muito tempo computacional se impusermos um limite de banda B aos nossos sinais, a complexidade computacional de (2.18) é $O(B^5)$. Essa complexidade pode ser reduzida para $O(B^4)$ se fizermos os cálculos no espaço harmônico. Para isso precisamos da decomposição de Fourier de f e h, que é dada por

$$f(\theta, \varphi) = \sum_{l=0}^{B-1} \sum_{m=-l}^{l} a_{lm} Y_m^l(\theta, \varphi) \quad \text{e} \quad h(\theta, \varphi) = \sum_{l=0}^{B-1} \sum_{m=-l}^{l} b_{lm} Y_m^l(\theta, \varphi). \tag{2.19}$$

Usando a notação (2.2) para economizar espaço e substituindo as equações acima em (2.18), podemos escrever a correlação da forma

$$\begin{aligned} C(g) &= \int_{S^2} \left[\sum_{l=0}^{B-1} \sum_{m=-l}^{l} a_{lm} Y_m^l(\omega) \right] \overline{\left[\Lambda(g) \sum_{l'=0}^{B-1} \sum_{m'=-l'}^{l'} b_{l'm'} Y_{m'}^{l'}(\omega) \right]} d\omega, \\ &= \sum_{l=0}^{B-1} \sum_{m=-l}^{l} \sum_{l'=0}^{B-1} \sum_{m'=-l'}^{l'} a_{lm} \overline{b_{l'm'}} \int_{S^2} Y_m^l \overline{\Lambda(g) Y_{m'}^{l'}}. \end{aligned} \tag{2.20}$$

Nesse ponto, podemos remover uma soma, notando que a integral só é diferente de zero para o caso $l = l'$. Dessa forma ficamos com

$$C(g) = \sum_{l=0}^{B-1} \sum_{m=-l}^{l} \sum_{m'=-l}^{l} a_{lm} \overline{b_{lm'}} \int_{S^2} Y_m^l \overline{\Lambda(g) Y_{m'}^{l}}. \tag{2.21}$$

Lembrando agora que a rotação de um harmônico esférico de grau l pode ser expressa como uma combinação linear de harmônicos esféricos de mesmo grau, como dado pela equação (2.9), temos a simplificação

$$\begin{aligned} C(g) &= \sum_{l=0}^{B-1} \sum_{m=-l}^{l} \sum_{m'=-l}^{l} a_{lm} \overline{b_{lm'}} \int_{S^2} Y_m^l(\omega) \overline{\sum_{k=-l}^{l} Y_l^k(\omega) D_{km'}^l(g) d\omega}, \\ &= \sum_{l=0}^{B-1} \sum_{m=-l}^{l} \sum_{m'=-l}^{l} \sum_{k=-l}^{l} a_{lm} \overline{b_{lm'} D_{km'}^l(g)} \int_{S^2} Y_l^m(\omega) \overline{Y_l^k(\omega)} d\omega. \end{aligned} \tag{2.22}$$

A integral (2.22) só é diferente de zero para o caso $k = m$, portanto temos finalmente

$$C(g) = \sum_{l=0}^{B-1} \sum_{m=-l}^{l} \sum_{m'=-l}^{l} a_{lm} \overline{b_{lm'} D_{mm'}^l(g)}. \tag{2.23}$$

Usando as relações de simetria dos coeficientes a_{lm} e das funções D_{mn}^l, podemos escrever a equação acima da forma

$$C(g) = \sum_{l=0}^{B-1} \sum_{m,n=-l}^{l} \overline{a_{lm}} b_{ln} D^{l}_{mn}(g). \tag{2.24}$$

Substituindo a equação 2.6 na equação acima ficamos com

$$C(g) = \sum_{l=0}^{B-1} \sum_{m,n=-l}^{l} \overline{a_{lm}} b_{ln} e^{-im\alpha} d^{l}_{mn}(\beta) e^{-in\gamma}. \tag{2.25}$$

A forma da equação (2.25) é especialmente adequada ao cálculo da correlação quando a malha escolhida tem as coordenadas α e β amostradas em pontos equidistantes, permitindo assim o uso da transformada rápida de Fourier. Se a coordenada β também puder ser amostrada em intervalos de mesmo tamanho, podemos estender o uso da transformada rápida de Fourier aos três ângulos de Euler. Substituindo a expansão de Fourier das funções $d^{l}_{mn}(\beta)$, dada pela equação A.9 em (2.25) temos

$$C(\alpha, \beta, \gamma) = \sum_{l=0}^{B-1} \sum_{m,n,u=-l}^{l} a_{lm} \overline{b_{ln}} B^{l}_{mnu} e^{-(im\alpha + iu\beta + in\gamma)}, \tag{2.26}$$

que foi a forma usada para implementação do nosso algoritmo de busca por multipletos.

2.5 Tesselação em S^2

Quando trabalhamos com funções definidas na esfera, somos obrigados a escolher uma malha na superfície da esfera (i.e. os pontos onde as funções serão amostradas). Esse problema é trivial em análise de imagens no plano, que, por não ter curvatura, permite o uso de uma malha cujos pixeis tenham áreas iguais. Quando tentamos estender o uso de malhas com pixeis de mesma área para a esfera, enfrentamos um problema. A complexidade computacional introduzida pode inviabilizar o cálculo de integrais na esfera e assim seu uso em um grande número de simulações.

Em geral, a viabilidade computacional está associada ao uso do algoritmo FFT no cálculo das integrais. Por exemplo, a malha definida em (2.14) permite o uso de FFT por usar intervalos constantes das coordenadas α, β e γ. Quando a malha não tem essa propriedade, temos que desistir do uso de FFT e usar relações de recorrência que, em geral, são mais lentas.

Vejamos algumas das malhas mais usadas.

Healpix. A divisão proposta pelos autores do software Healpix, que significa *Hierarchical Equal Area isoLatitude Pixelization* não é exata, mas pode ser muito precisa conforme aumentamos o número de pixeis. Sua característica mais desejável é que os pixeis, apesar de terem formas diferentes uns dos outros, têm a mesma área. Essa característica é muito importante por dois motivos. O primeiro é que a todo experimento está associada uma resolução angular. No caso do Observatório Pierre Auger por exemplo, para energias maiores que 5 EeV podemos considerar que a resolução angular é menor que 1°. Isso quer dizer que se pudermos fazer todos os pixeis com aproximadamente este tamanho, teremos um melhor aproveitamento de nossa divisão. Além disso, uma vez que os pixeis têm áreas iguais, qualquer ruído que seja uniforme ao longo da superfície, contribui igualmente para todos os pixeis, fato que não ocorreria se as áreas não fossem iguais.

Apesar de conceitualmente ser mais fácil tratar os dados usando uma divisão em áreas iguais, seu beneficio real não está claro, já que nenhum experimento tem uma exposição igual em todas as direções, sendo necessário atribuir pesos diferentes a pixeis com exposições diferentes. No caso do Observatório Pierre Auger, os detectores estão expostos apenas à metade do céu.

A divisão Healpix apresenta um desafio computacional pois ela não possui uma divisão igualmente espaçada da coordenada θ, que impossibilita o uso da transformada de Fourier para calcular as integrais. O uso de relações de recorrência deve ser feito para cada ponto amostrado da coordenada θ.

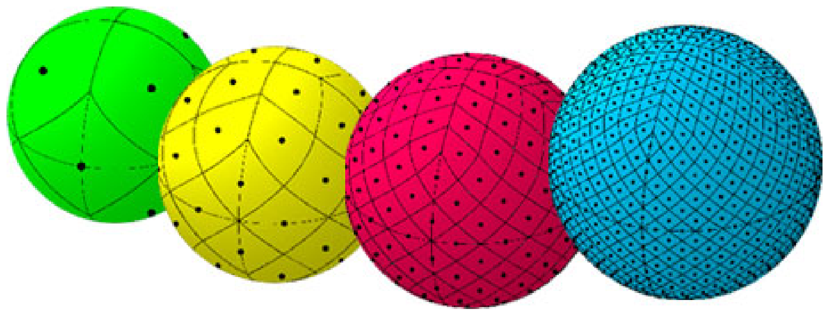
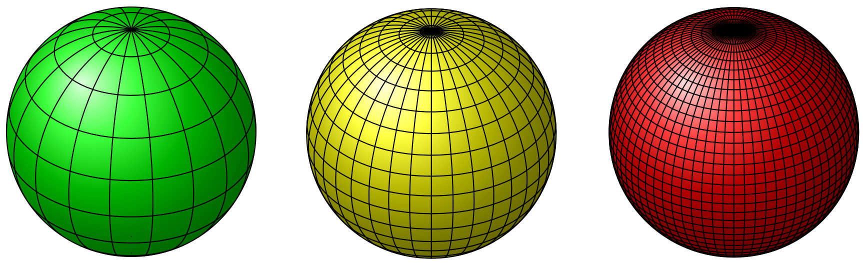


Figura 2.2: Tesselação Healpix para diferentes números de Pixel.

Projeção cilíndrica equidistante (ECP). A projeção ECP é a mais simples de todas do ponto de vista computacional. Nessa tesselação temos tanto em θ quanto em φ divisões equidistantes. O que permite o cálculo das integrais via FFT. Além disso, é muito fácil diminuir a resolução de um mapa combinando pixeis vizinhos, procedimento que é a base das divisões do tipo *igloo*.

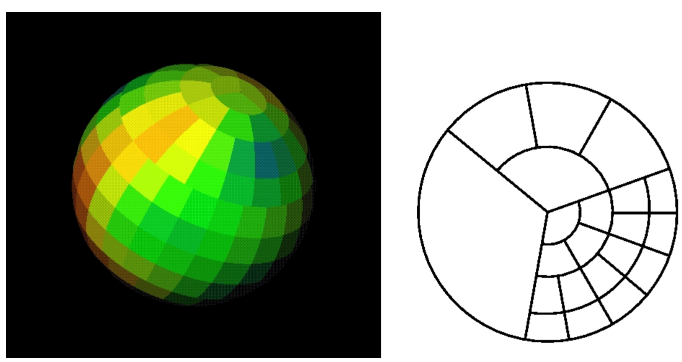
O grande problema dessa tesselação é que os pixeis que estão perto dos polos são distorcidos e se tornam muito pequenos, chegando ao ponto de serem possivelmente menores que a própria resolução angular do experimento. Isso significa que, em uma análise de dados, esses pixeis serão praticamente dominados por ruído ou não conterão nenhum evento.

Divisão do tipo iglu. Divisões do tipo iglu oferecem tanto pixeis de áreas iguais quanto transformadas rápidas na esfera, mas requerem uma técnica relativamente sofisticada para agrupar pixeis de mesma latitude Por exemplo, partindo da tesselação ECP, podemos agrupar pixeis de tal forma que a área deles não varie conforme nos aproximamos do zênite.



(a) (b) (c)

Figura 2.3: Tesselação ECP (Equidistant cylindrical projection)



(a) (b)

Figura 2.4: Tesselação do tipo *igloo*.

Capítulo 3

Wavelets esféricos

Everything should be made as simple as possible, but not simpler.
ALBERT EINSTEIN

People often choose something new because they don't know its weaknesses yet.
B. STROUSTRUP

A ANÁLISE WAVELETS é um assunto novo em comparação com a bem estabelecida análise de Fourier e encontra sua utilidade quando o caráter global de senos e cossenos deixa de ser a forma ideal de decompor um sinal. Isso ocorre, por exemplo, quando estudamos transições abruptas, em que a decomposição em senos e cossenos precisa de um número grande de coeficientes para descrever o sinal com precisão.

Quando analisamos um sinal, seja ele uma série temporal ou uma imagem, por exemplo, podemos pensar em duas maneiras de tratar o problema. Primeiramente podemos trabalhar diretamente com os valores dos pixeis, como no caso de uma imagem. Apesar de termos acesso a toda a informação contida no sinal dessa forma, ela se torna inadequada quando queremos informações mais gerais; por exemplo, em uma série temporal o valor de um pixel nos dá informação para apenas um determinado instante de tempo. No outro extremo, podemos decompor o sinal em funções conhecidas, como senos e cossenos, que têm valores diferentes de zero em todo o domínio de análise e portanto fornecem informações globais i.e. quais as frequências mais importantes contidas no sinal.

Muitas vezes no entanto, nenhuma dessas formas é a ideal. No caso em que temos uma região de descontinuidade, a análise de Fourier é inadequada pois precisa de muitos coeficientes para descrever as características importantes da descontinuidade i.e. sua localização, por exemplo. Nesse caso precisamos de uma decomposição mista, que não seja nem tão localizada quanto a análise pixel a pixel, nem tão global quanto a análise de Fourier. Aqui surge a utilidade da análise de wavelets, através da qual conseguimos concentrar as informações relevantes de um sinal em alguns poucos coeficientes.

Nesse trabalho desenvolvemos um algoritmo baseado em wavelets esféricos para a detecção de estruturas filamentares que se formam em mapas de raios cósmicos. Esse capítulo será usado para dar detalhes e exemplos de wavelets. No próximo capítulo, falaremos sobre o algoritmo que usamos para identificar essas estruturas filamentares, chamadas de multipletos.

3.1 Introdução

No capítulo 2 vimos que, da mesma forma que uma série temporal pode ser decomposta em uma série de senos e cossenos, uma função $f : S^2 \to \mathbb{R}$, pode ser decomposta em uma soma de harmônicos esféricos, que são denotados por $Y^l_m(\theta, \varphi)$, como mostrado na equação (2.10). Os coeficientes a_{lm} da expansão são chamados de multipolos ou simplesmente, coeficientes de Fourier de f. No formalismo que adotamos para a análise de wavelets, os coeficientes a_{lm} são fatorados da seguinte forma

$$a_{lm} = k(l) S_{lm}. \tag{3.1}$$

com $l \in \mathbb{N}$ e $|m| \leq l$.

Nesse formalismo expansões (ver adiante) são aplicadas apenas ao *kernel* $k(l)$, mantendo os coeficientes S_{lm}, chamados de coeficientes de direcionabilidade, intactos.

Os coeficientes S_{lm}, por sua vez, controlam a capacidade dos wavelets de encontrar direções de sinais na esfera. Eles satisfazem as mesmas relações de simetria que os coeficientes a_{lm}. Adicionalmente impõe-se o vínculo

$$\sum_{m=-l}^{l} S_{lm}^2 = 1, \tag{3.2}$$

que garante a normalização dos wavelets.

A expansão dos wavelets por um parâmetro de escala $a \in (1, \infty)$ é definida na forma

$$(f_a)_{lm} = k(al) S_{lm}. \tag{3.3}$$

Para uma função com limite de banda B como definido em (2.12) vemos que a estará contido no intervalo $a \in [0, B]$, uma vez que fora desse intervalo $f_{lm} = 0$.

Nas seções seguintes vamos mostrar a definição do kernel $k(l)$ e dos coeficientes S_{lm} usados nesse trabalho.

3.2 O kernel $k(l)$

O KERNEL introduzido na equação (3.3) é obtido primeiramente de um kernel contínuo através de integrações por fatias do parâmetro de escala a. O primeiro passo para calculá-lo se baseia no cálculo da função de escala $\Phi(k)$ a partir da equação

$$\Phi^2(l) = \int_1^{\infty} \frac{da}{a} \hat{K}(al), \tag{3.4}$$

onde $K(al)$ é o kernel contínuo definido por

$$\hat{K}(l) = \begin{cases} exp\left(-\frac{1}{1-t^2(l)}\right) & \text{se } t(l) \in (-1,1) \\ 0 & \text{se } t(l) \notin (-1,1) \end{cases} \tag{3.5}$$

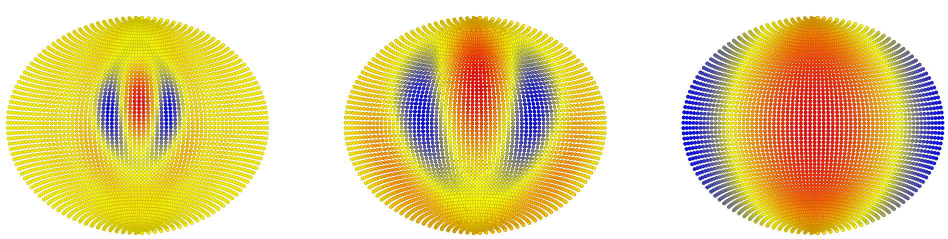
onde a função $t(l)$ é dada por

$$t(l) = 2\frac{al - B}{(a-1)B} - 1. \tag{3.6}$$

No caso do kernel discreto $k(l)$ assumimos que o parâmetro de escala a tem a forma $a = 2^j$ com $j \in \mathbb{N}$ onde j passa a ser nosso parâmetro de escala discreta. A maior escala possível será denotada por J, assim $0 \leq j \leq J$. Com isso, os valores de $k(l)$ na escala j podem ser definidos, e segue

$$k^2(l) = \Phi^2(2^{-j}l) - \Phi^2(l). \tag{3.7}$$

Na figura 3.1 vemos um exemplo de expansão de um wavelet. Nesse exemplo mantivemos $J = 5$ e variamos j entre 3, 4 e 5. Podemos ver que o tamanho angular do wavelet aumenta com a escala. A sensibilidade de um wavelet em identificar um sinal é maior no caso em que o sinal e o wavelet tem aproximadamente o mesmo tamanho angular.



(a) (b) (c)

Figura 3.1: Exemplo da expansão de um wavelet com parâmetros $J = 5$, $N = 3$ e $j = 3, 4, 5$ respectivamente.

Tabela 3.1: Primeira coluna: escala j. Segunda coluna: suporte do wavelet. Terceira coluna: Tamanho angular em graus. Quarta coluna: precisão angular máxima nas variáveis angulares.

j	Suporte	Tamanho angular	Precisão
0	$(256, 128)$	$(0.7, 1.4)$	0.7
1	$(256, 64)$	$(0.7, 2.8)$	0.7
2	$(128, 32)$	$(1.4, 5.6)$	1.4
3	$(64, 16)$	$(2.8, 11.3)$	2.8
4	$(32, 8)$	$(5.6, 22.5)$	5.6
5	$(16, 4)$	$(11.3, 45.0)$	11.3
6	$(8, 2)$	$(22.5, 90.0)$	22.5
7	$(4, 1)$	$(45.0, 180.0)$	45.0
8	$(2, 1)$	$(90.0, 180.0)$	90.0

Suporte harmônico compacto. A ideia de suporte harmônico compacto é muito importante para o processamento do sinal. Dizemos que uma função f tem o suporte harmônico compacto no intervalo $l \in (\alpha^{-1}B, B)$, se

$$f_{lm} = 0 \quad \forall \quad l, m \quad \text{com} \quad l \notin (\alpha^{-1}B, B). \tag{3.8}$$

Dessa forma, para obter uma função f com o suporte compacto, basta escolhermos um kernel $K(l)$ que também tenha o suporte compacto nesse mesmo intervalo. Na figura 3.2 podemos ver o suporte dos wavelets definidos na seção 3.2.

3.3 Coeficientes S_{lm}

A definição dos coeficientes S_{lm} é de fundamental importância para este trabalho. Os coeficientes a_{lm} resultantes da decomposição de um sinal em harmônicos esféricos têm em geral valores diferentes de zero sempre que $l < B$ e $|m| \leq l$, onde B é o limite de banda do sinal, acima do qual todos os coeficientes têm valor zero. Para viabilizar computacionalmente o cálculo da

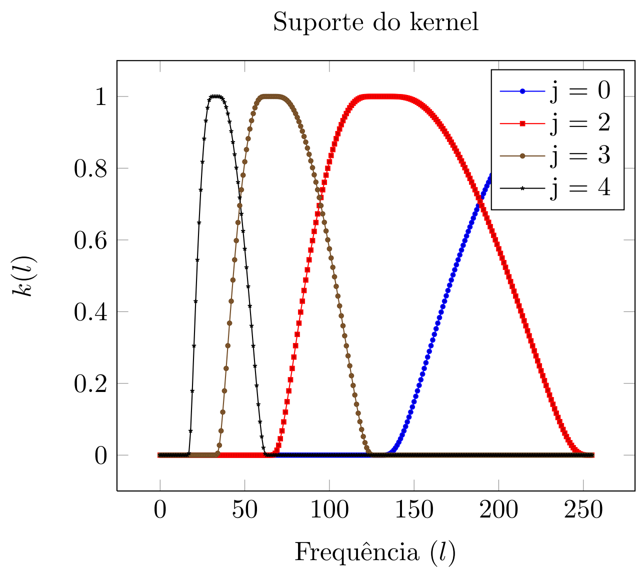


Figura 3.2: Suporte do kernel para várias escalas j e $J = 8$.

correlação (2.24) temos que diminuir a precisão do sinal. Para isso é necessário introduzir o conceito de direcionabilidade, que está ligado à capacidade de um wavelet encontrar orientações de estruturas na esfera.

Definimos uma função $g(\omega) \in L^2(S^2)$ como direcionável, se qualquer rotação dela em torno de si mesma pode ser escrita como uma combinação linear de um número finito M de funções base g_i

$$f_\zeta(\omega) = \sum_{i=0}^{M-1} k_i(\zeta) g_i(\omega). \tag{3.9}$$

As funções $k_i(\zeta)$ com $0 \leq \zeta < \pi$ são funções peso. A definição acima implica um limite de banda azimutal para o sinal, que pode ser expresso da forma

$$g_{lm} = 0 \quad \forall \quad l, m \quad \text{com} \quad |m| \geq N. \tag{3.10}$$

Em outras palavras, não precisamos considerar todas as rotações possíveis no cálculo da correlação, e basta que sejam feitos apenas para N rotações.

Os wavelets que estamos usando são direcionáveis e têm sua forma dada por

$$S_{lm} = \eta_N \beta_{(N,m)} \left[\frac{1}{2^{\gamma(N,l)}} \binom{\gamma(N,l)}{\frac{\gamma(N,l)-m}{2}} \right]^{1/2}, \tag{3.11}$$

com

$$\eta_N = \begin{cases} i & \text{se N par} \\ 1 & \text{se N ímpar} \end{cases}, \tag{3.12}$$

e

$$\beta_{(N,m)} = \begin{cases} 0 & \text{se N + m par} \\ 1 & \text{se N + m ímpar.} \end{cases}, \tag{3.13}$$

$$\gamma(N,l) \;\; = \;\; min\left(N-1, l - \frac{1+(-1)^{N+l}}{2}\right). \tag{3.14}$$

Na figura abaixo vemos o significado físico da direcionalidade. A função mais à direita é a soma das duas funções à sua esquerda.

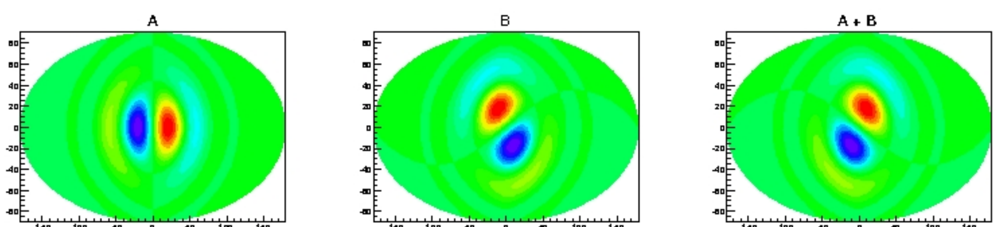
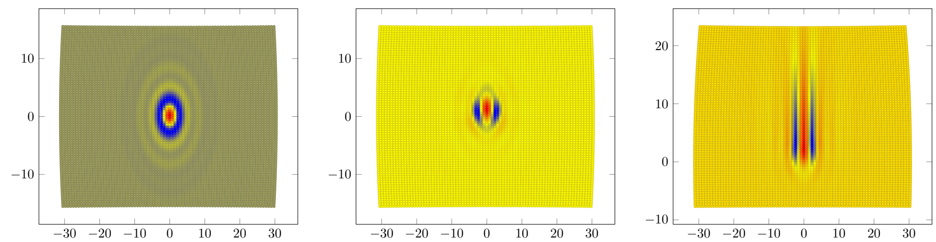


Figura 3.3: Controle de direcionabilidade. A função mais à direita é a soma das duas funções à sua esquerda

Analogamente ao que fizemos na figura 3.1, podemos manter J e a escala j constantes e variar apenas o parâmetro N. Por exemplo, se pretendemos



(a) (b) (c)

Figura 3.4: Exemplo do controle da direcionabilidade de um wavelet a partir do parâmetro N. Usamos $J = 8$, $j = 2$ e $N = 1, 3, 127$. O wavelet mostrado na figura (c) também foi usado em nossos algoritmos para encontrar multipletos. Os eixos usam unidades de graus.

identificar estruturas como fontes pontuais, que têm simetria em torno de um eixo, podemos escolher $N = 1$, pois não há direções preferidas nesse caso. Já no caso de uma estrutura filamentar, gostaríamos de ter a maior precisão possível, que nesse caso é dada por $N = 2^{J-j+1}$. Nesse trabalho por exemplo, usamos $(J, j, N) = (8, 1, 127)$ na busca por multipletos uma vez que N = 127 é o maior valor possível de N para $j = 1$ e $J = 8$.

Se o tempo computacional for muito alto para um dado N podemos escolher valores menores, perdendo um pouco de precisão, mas sem a necessidade de mudar a escala por exemplo.

Na figura 3.4, vemos um exemplo em que N assume os valores 1, 3 e 127 com $J = 8$ e $j = 2$.

3.4 Reconstrução do sinal

Até agora, vimos como obter a representação de wavelets de um sinal $f(\theta, \varphi)$, dado pela equação (2.26). Em muitas situações, no entanto, precisamos fazer o caminho inverso, ou seja, dada a representação do sinal f no espaço de wavelets, denotado por $C(\alpha, \beta, \gamma)$, queremos recuperar f. Esse procedimento tem aplicações importantes em tratamento de sinais, pois permite que operações como filtragem do sinal e tratamentos estatístico, sejam aplicadas no espaço de wavelets, onde certas características do sinal são mais evidentes.

Denotando os multipolos do sinal f por a_{lm} e os multipolos do wavelet por $(b^j)_{lm}$, onde j representa a escala do wavelet, temos

$$a_{lm} = \delta_{0l}\delta_{0m}a_{00} + \frac{2l+1}{8\pi^2}\sum_{j=0}^{J}\sum_{n=-l}^{n=l}(b^j)_{lm}(f^j)^l_{mn}, \tag{3.15}$$

onde os coeficientes $(f^j)^l_{mn}$ são dados pela equação (2.16). Uma vez que os coeficientes a_{lm} tenham sido obtidos, uma simples transformada em harmônicos esféricos, dada pela equação (2.10), pode recuperar o sinal f.

Capítulo 4

Procurando por multipletos

What we have to learn to do we learn by doing.
— ARISTOTLE, Ethica Nicomachea II (c. 325 B.C.)

NESTE CAPÍTULO mostramos com detalhes o algoritmo desenvolvido para a busca de multipletos nos dados do Observatório Pierre Auger. A busca por multipletos é um tópico muito ativo na Colaboração Pierre Auger devido a sua importância no estudo de fontes de raios cósmicos. Como o Observatório Pierre Auger está contribuindo de forma decisiva para a amostragem na região de altas energias, surge pela primeira vez a possibilidade clara de identificação de fontes com estatística significativa.

O número de trabalhos relacionados a identificação de multipletos ainda é muito discreto quando comparado à maioria dos temas relacionados à física de raios cósmicos. Por exemplo, oficialmente a Colaboração Pierre Auger tem apenas um trabalho [7] publicado sobre o tema; os artigos em forma de *Gap Notes*, que são Notas Internas privadas aos membros da Colaboração, fornecem outro meio para comunicação sobre o assunto. Até o momento onde escrevo apenas 10 *Gap Notes* foram publicadas sobre o tema. Fora do Observatório Pierre Auger, o autor dessa tese encontrou apenas um outro artigo que trata do problema [16]. Dessa forma, vemos que apesar da identificação de multipletos ter consequências físicas muito importantes, existem poucos cientistas trabalhando no problema.

Na seção 1.4 vimos a definição de multipletos e uma revisão dos trabalhos mais importantes, que tratam de sua identificação. Nesse capítulo apresentamos os detalhes do algoritmo desenvolvido. Mas antes disso veremos com mais detalhes os desafios que enfrentaremos.

4.1 Busca por força bruta

Qual o algoritmo mais simples que podemos imaginar para procurar multipletos em um conjunto de dados? Se os critérios que caracterizam um conjunto de eventos como um multipleto estiverem bem definidos, como, por exemplo, o alinhamento dos eventos e sua ordenação energética, tudo que temos a fazer é testar todas as combinações possíveis e selecionar os conjuntos que satisfazem os critérios. Vejamos a viabilidade computacional desse procedimento e de suas versões mais elaboradas.

Primeira tentativa. A forma mais óbvia de procurar por multipletos é pelo uso da força bruta, ou seja, testar quais conjuntos de k eventos em um fundo de n eventos, satisfazem os critérios para serem considerados candidatos. O número total de combinações nesse caso é dado pelo coeficiente binomial $\binom{n}{k}$.

Para fazer uma estimativa desse número vamos usar valores compatíveis com o Observatório Pierre Auger, que é atualmente o experimento que mais coleta dados na faixa de energia de nosso interesse. O número total de eventos no intervalo de energia $E > 15$ EeV é de aproximadamente 4247, portanto para fins de estimativa usaremos $n = 5000$.

É difícil estimar o número k de eventos que compõem um multipleto. A publicação da Colaboração Pierre Auger sugere que candidatos com menos de $k = 10$ não apresentam estatística significativa. Portanto usaremos esse número em nossas estimativas. O número total de conjuntos que temos que analisar é portanto $\binom{n}{k} = n!/k!(n-k)!$. No nosso caso os valores de $n!$

e $(n-k)!$ são muito altos e portanto justifica-se o uso da aproximação de *Stirling* $n! \approx \sqrt{2\pi n}(n/e)^n$. O erro dessa aproximação é de $1/12n$, ou seja, da ordem de $\mathrm{O}(10^{-5})$ para os números envolvidos. Temos portanto

$$\begin{aligned}
\binom{n}{k} &= \frac{n!}{k!(n-k)!}, \\
&\approx \frac{\sqrt{2\pi n}(n/e)^n}{k!\sqrt{2\pi(n-k)}((n-k)/e)^{n-k}}, \\
&\approx \frac{(n/e)^k(1-k/n)^{k-n-1/2}}{k!}, \\
&\approx \frac{(n/e)^k(1-(k-n-1/2)k/n)}{k!}, \qquad (4.1)
\end{aligned}$$

onde na última aproximação, usamos a expansão binomial para o cálculo de $(1-k/n)^{k-n-1/2}$. Dessa forma concluímos que o número total de possibilidades é de $\binom{5000}{10} \approx 1,34 \times 10^{26}$.

Devido à magnitude desse número, um algoritmo que testa todas as possibilidades é inviável computacionalmente. No entanto, estudando cuidadosamente o problema, vemos que nem todas as combinações precisam ser testadas, se assumirmos alguns fatos sobre multipletos.

Segunda tentativa. No capítulo 1, vimos que a deflexão sofrida por partículas com energias muito altas é de apenas alguns graus. Dessa forma sabemos que a maior parte das combinações calculadas em nossa primeira tentativa, não resultarão em multipletos e podem ser ignoradas. Precisamos analisar apenas os conjuntos de eventos cuja diferença angular entre o evento de maior e menor energia, não ultrapasse um certo valor de corte.

Um algoritmo possível seria: definir uma malha na superfície da esfera, como apresentado na seção 2.5. A cada ponto da malha associamos um círculo com um certo diâmetro e selecionamos os eventos que incidem nesse círculo. Calculando todas as possibilidades de multipletos nesses conjuntos reduzidos de eventos, teríamos todas as possibilidades testadas. Obviamente esse procedimento introduz um erro devido ao espaçamento entre os pontos da malha, que pode ser feito tão pequeno quanto se queira, aumentando a resolução da malha (resultando em maior tempo computacional de análise).

Para estimar o número de possibilidades nesse procedimento vamos assumir que temos um conjunto de eventos distribuídos uniformemente no céu e que a malha usada é dada pela tesselação ECP, apresentada na seção 2.5. O número total de pontos dessa tesselação na discretização que adotamos é dado por $n = 4B^2$, onde B, chamado de *limite de banda* do sinal, está relacionado com o espaçamento entre os pontos da malha. A área onde incidem os eventos associados a cada círculo pode ser calculada facilmente para o ponto localizado no polo da esfera

$$\begin{aligned} A_c &= \int_0^\theta \int_0^{2\pi} r^2 \sin\theta d\theta d\varphi, && (4.2) \\ &= 2\pi r^2(1-\cos\theta). && (4.3) \end{aligned}$$

O número total de eventos n_c incidentes em um círculo é dado por $n_c = n_t A_c/A_e$, onde n_t é o número total de eventos no céu isotrópico, A_c a área do círculo e A_e a área da esfera.

Dessa forma temos $n_c = n_t(1-\cos\theta)/2$. Para $n_t = 5000$ e $\theta = 10°$, temos $n_c \approx 38$ eventos. Dessa forma o número total de combinações associado a um ponto da malha é em média $\binom{38}{10} = 472.733.756$. Multiplicando esse valor pelo número total de pontos da malha, temos o número total de operações a serem feitas

$$\begin{aligned} 4B^2\binom{38}{10} &= 4(256)^2 472.733.756, \\ &= 123.924.317.732.864, \\ &\approx 1,23 \times 10^{14}, && (4.4) \\ & && (4.5) \end{aligned}$$

onde usamos $B = 256$. Esse valor de B produz um espaçamento de aproximadamente $0.7°$ para pontos próximos ao equador. Apesar de $1,23 \times 10^{14}$ ser muito menor do que $1,34 \times 10^{26}$, ainda não é viável computacionalmente, principalmente no caso em que muitas simulações devem ser feitas.

Nesse procedimento o tempo gasto para gerar as combinações de eventos e testá-las não é o único a ser levado em conta no tempo computacional total. Temos um novo fator importante. Para selecionar os eventos incidentes em cada círculo, temos que percorrer todos os dados novamente, ou seja, vamos percorrer os $n = 5000$ eventos $4B^2 = 262144$ vezes, fato que não ocorreu em nossa primeira tentativa.

Como antes, ainda existem fatos sobre multipletos que podem ser usados para simplificar o algoritmo. Esse é o ponto de nossa terceira tentativa.

Terceira tentativa. Em vez de testarmos todas as combinações dentro dos círculos como fizemos em nossa segunda tentativa, podemos assumir que os eventos associados a um multipleto estão distribuídos ao longo do segmento de uma geodésica que passa pelo ponto central do círculo (as geodésicas na superfície da esfera são curvas formadas pela intersecção de um plano que passa pelo centro da esfera com sua superfície, também chamados de *grandes círculos*), dessa forma, em vez de testar todas as possibilidades dentro do círculo precisamos analisar apenas os eventos distribuídos ao longo desses segmentos, como mostrado na figura 4.1. Obviamente temos uma quantidade infinita de orientações a testar, mas se o erro introduzido por um discretização da variável γ, que descreve o orientação do segmento for aceitável, temos um procedimento viável computacionalmente.

Se usarmos o mesmo limite de banda B para γ como foi usado para α e β, o número total de conjuntos de eventos que temos que testar é dado por

$(2B)^3$. Usando $B = 256$ temos

$$\begin{aligned} (2B)^3 &= 2^3(2^8)^3, \\ &= 2^{27}, \\ &= 134.217.728, \\ &\approx 1,34 \times 10^8. \end{aligned}$$

Lembramos aqui, que para formar cada um dos $(2B)^3$ conjunto de eventos usados nessa análise, precisamos percorrer todos os dados e selecionar apenas os eventos que incidem na tira.

Apesar da aparente viabilidade computacional, esse algoritmo deixa de ser atraente pelo fato de ser muito sensível a ruído. Quando selecionamos os eventos incidentes em um tira, não temos a garantia de que incidirão apenas eventos pertencentes a um multipleto. Eventos pertencentes ao fundo também estarão presentes, gerando ruído. Veremos que o observável que usamos para classificar os multipletos como candidatos é muito sensível a ruído, e portanto o algoritmo não é adequado como parâmetro de seleção de candidatos a multipleto.

4.2 Novo algoritmo

Descrevemos abaixo o algoritmo que desenvolvemos para busca de multipletos. Esse algoritmo se baseia no uso de wavelets esféricos para encontrar a localização α e β e orientação γ dos multipletos na esfera celeste. Os dados de entrada são as direções de incidência (θ, φ) dos raios cósmicos juntamente com sua energia E.

Como configuração da análise, escolhemos qual wavelet será usado. A escolha adequada depende do problema em mão e é feita ajustando os parâmetros J, j e N como descrito no capítulo 3. Uma vez feita essa escolha, o algoritmo segue da forma

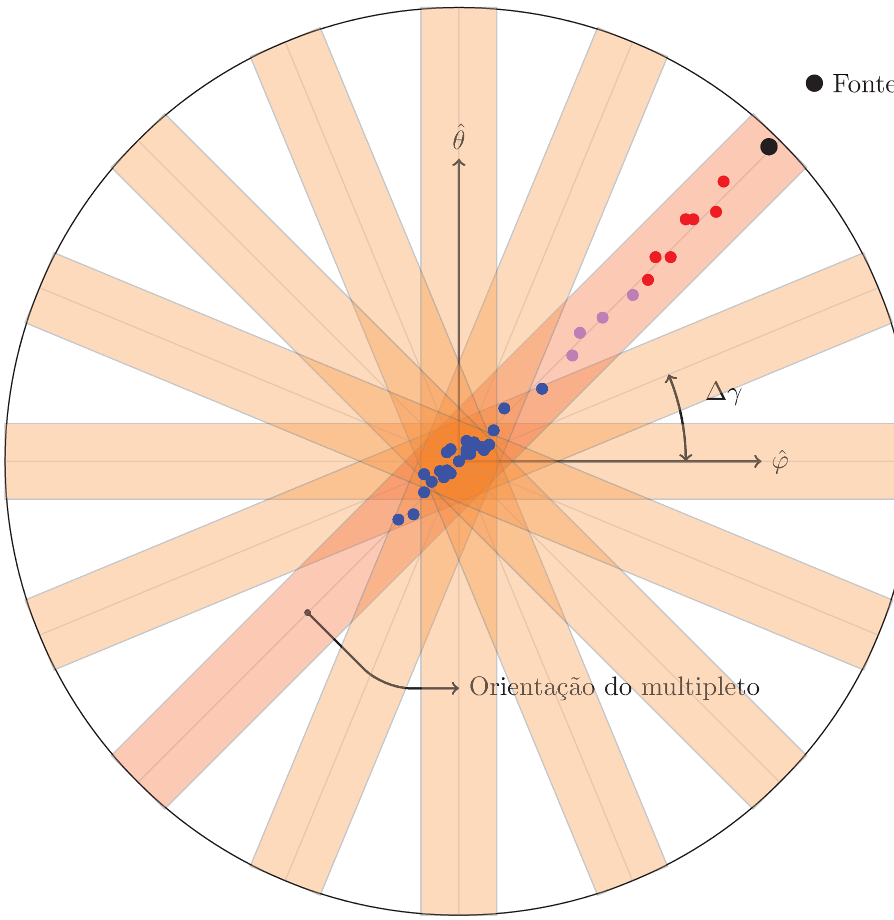


Figura 4.1: Variando a orientação das tiras podemos cobrir toda a área do círculo.

1. Para um dado conjunto de eventos, calcular a função $f(\theta, \varphi)$, que representa o sinal. Isso envolve escolher uma tesselação na esfera e contar o número de eventos que atingem cada pixel da esfera.

2. Calcular a expansão de Fourier da função $f(\theta, \varphi)$, resultando nos coeficientes a_{lm}, como descrito na equação (2.10).

3. Calcular a representação de wavelets do sinal, dada pela equação (2.26), com b_{lm} dado pela equação (3.1).

4. Selecionar as triplas $(\alpha, \beta, \gamma)_i$, para as quais a correlação $C(\alpha, \beta, \varphi)_i > C_0$, onde C_0 é um valor de corte, resultando em M coeficientes C.

5. Selecionar os eventos que atingem o céu em cada (α, β, γ) selecionado, resultando em M grupos de eventos.

6. Descartar todos os grupos de eventos para os quais $n < n_0$, onde n é o número de eventos no grupo e n_0 é um valor de corte.

7. Para cada grupo de eventos restantes, calcular a correlação c do gráfico $E \times \delta$, onde E é a energia do evento e δ o ângulo de deflexão. A deflexão é calculada em relação a um dos eventos do grupo.

8. Aceitar como candidato a multipleto, todos os grupos de eventos, para os quais $c > c_0$, onde c_0 é um valor de corte.

Estabelecendo valores de corte. O algoritmo apresentado acima pressupõe três valores de corte, C_0, n_0 e c_0. Quanto melhor pudermos estimar estes valores, melhor será a confiança nos resultados obtidos, por isso sugerimos aqui como isso pode ser feito.

Primeiramente precisamos conhecer os dados onde queremos buscar multipletos. As duas quantidades de interesse aqui são a exposição do experimento, isto é, quantos eventos em uma dada direção são esperados e a distribuição de energia que eles seguem. Uma vez que tenhamos essas informações podemos simular M conjuntos de dados com as mesmas características, mas onde não há campo magnético, ou seja onde um multipleto só pode ocorrer ao acaso. Dessa forma seremos capazes de saber com que frequência multipletos ocorrem em um céu isotrópico, ou em outras palavras, poderemos responder com que frequência teremos $|C| > C_0$, $n > n_0$ e $c > c_c$, ocorrendo ao acaso.

Para determinar os valores podemos usar o algoritmos da seguinte forma: Aplicar o algoritmo em M céus simulados, onde no passo 5 escolhemos apenas o maior coeficiente C. Com isso chegaremos ao passo 7 com apenas uma correlação c. Com M valores de C e c, calculamos os valores de corte da forma

$$C_0 = \bar{C} + a\sigma_C \quad \text{e} \quad c_0 = \bar{c} + b\sigma_c, \tag{4.6}$$

onde $\bar{C}$ e $\bar{c}$ representam médias sobre M valores e σ_C e σ_c os respectivos desvios padrão. Os valores de a e b podem ser ajustados para tornar o critério de corte mais severo. Nesse trabalho usamos $a = b = 1$.

O valor de corte n_0 deve ser escolhido caso a caso. Por exemplo em nossa análise, simulamos um multipleto com dez eventos, e portanto consideramos $n_0 = 10$, ou seja, todos os candidatos a multipletos devem ter mais de dez eventos.

Seleção eventos. A distância infinitesimal ds entre dois pontos em coordenadas esféricas é dada por

$$ds^2 = dr^2 + r^2((d\beta)^2 + \sin^2\beta(d\alpha)^2). \tag{4.7}$$

Como estamos calculando distâncias na superfície de uma esfera, temos $dr = 0$. Com essa simplificação ds pode ser escrito da forma

$$ds^2 = r^2((d\beta)^2 + \sin^2\beta(d\alpha)^2). \tag{4.8}$$

Um vetor infinitesimal qualquer no plano tangente à esfera com origem em

(r, α, β) pode ser escrito da forma

$$\vec{d} = rd\beta\hat{\beta} + r\sin\beta d\alpha\hat{\alpha} \tag{4.9}$$

Agora escrevemos esse vetor em termos dos versores $\hat{L}$, que está alinhado com o multipleto e $\hat{W}$, que é perpendicular ao multipleto. Esses dois versores se relacionam com $\hat{\alpha}$ e $\hat{\beta}$ através de uma rotação de um ângulo γ

$$\hat{\alpha} = \cos\gamma\hat{L} - \sin\gamma\hat{W} \tag{4.10}$$
$$\hat{\beta} = \sin\gamma\hat{L} + \cos\gamma\hat{W}. \tag{4.11}$$

Com isso podemos escrever 4.9 da forma

$$\vec{d} = d_{\hat{L}}\hat{L} + d_{\hat{W}}\hat{W} \tag{4.12}$$

onde os deslocamentos nas direções $\hat{L}$ e $\hat{W}$ são dados por

$$d_{\hat{L}} = r(d\beta\cos\gamma + \sin\beta d\alpha\sin\gamma) \tag{4.13}$$
$$d_{\hat{W}} = r(-d\beta\sin\gamma + \sin\beta d\alpha\cos\gamma). \tag{4.14}$$

As equações acima foram usadas em nossa implementação para determinar quais eventos estão dentro de um segmento e quais não.

4.3 Análise em simulações com CRPropa

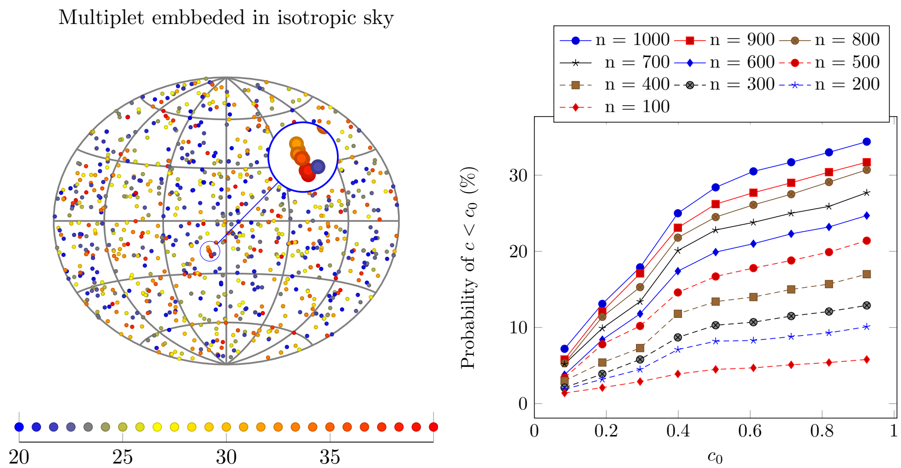
Antes de aplicar o algoritmo em dados reais do Observatório Pierre Auger, vamos testá-lo em dados simulados, onde o comportamento do algoritmo pode ser entendido com mais clareza.

Usamos o programa CRPropa para simular multipletos em um cenário controlado, onde as fontes injetam partículas com um espectro de energias que obedece a lei de potência $E^{-2.2}$. As fontes são consideradas pontuais e distam 30 Mpc do observador. O observador por sua vez é uma esfera com raio 1Mpc. As partículas propagam em um campo magnético de 1 nG.

Cada fonte injeta muitas partículas mas apenas algumas chegam ao observador, devido à deflexão sofrida pelo campo magnético. Devido a simetria esférica do problema, não temos a necessidade de variar a posição do multipleto, portanto, dos vários multipletos simulados, escolhemos apenas um para ser usado em nossa análise. O multipleto escolhido é composto de dez partículas, onde as partículas tem energia no intervalo $15 < E < 55$ EeV e sua correlação é de $c = 0.99$, ou seja, o gráfico $\delta \times 1/E$ se aproxima de uma linha reta.

A análise segue com os passos. Simulamos céus isotrópicos com os seguintes números de partículas $M = 100, 200, ..., 1000$. Para cada valor de M, fizemos 1000 simulações. Os eventos são simulados com exposição uniforme no céu e com ausência de campos magnéticos para defletir as partículas. Nesse caso um multipleto pode ocorrer somente ao acaso. A partir desse conjunto de dados, calculamos os valores de corte C_0 e c_0 para cada M, como explicado na seção 4.2.

Na segunda parte da análise usamos exatamente os mesmos eventos, com a exceção de que um multipleto gerado com o CRPropa é inserido nos dados. O objetivo aqui é ver se os wavelets são sensíveis o suficiente para detectar o multipleto imerso no fundo isotrópico. A fração de partículas varia de $f = 10^{-1}$ a $f = 10^{-2}$. Os gráficos comparando os dois casos podem ser vistos nas figuras 4.3. Na figura 4.3 (a) vemos que a correlação C segue bastante destacada para os dois casos até que a fração entre o número de eventos no multipleto e no fundo alcança $f = 10^{-2}$. Nesse ponto as barras de erros começam a se tocar. Portanto consideramos esse o limite de sensibilidade do algoritmo proposto. Na figura 4.3 (b) vemos a correlação c para os dois casos. Vemos claramente para o gráfico em vermelho que o valor de c decresce à medida que adicionamos mais eventos ao fundo. É notável que o algoritmo proposto não tenha identificado nenhum multipleto por engano (gráfico em



(a) Mapa com fração 10^{-2} (b) Erros tipo II

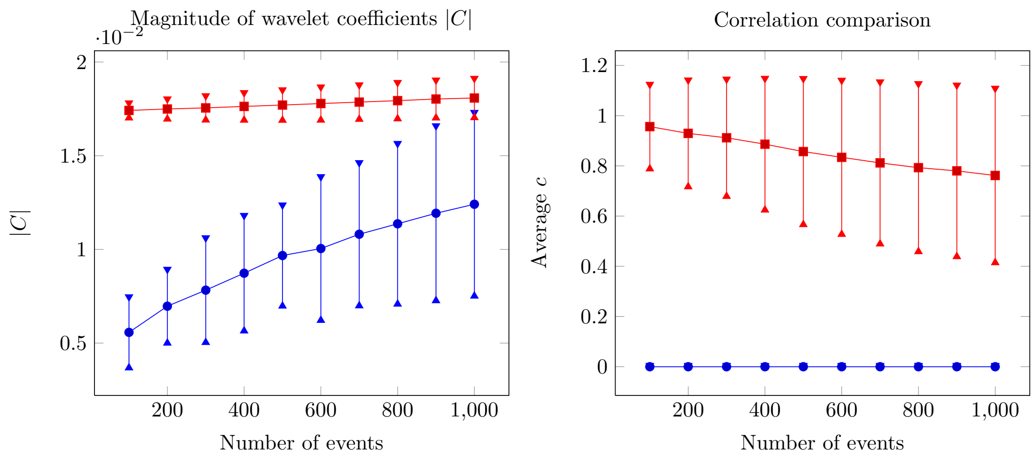
Figura 4.2: Na esquerda: Multipleto com dez eventos, imerso em um fundo com 1000 eventos. Na direita: Probabilidade do algoritmo encontrar $c < c_0$ para um multipleto cuja correlação é de $c = 0.99$.

azul). Atribuímos esse fato aos critérios severos que usamos, uma vez que candidatos a multipletos têm que passar por três critérios de corte i.e. C_0, c_0 e n_0.

Na figura 4.2 (b) vemos o erro de tipo II para esse algoritmo. Ou seja, a probabilidade de descartarmos um multipleto que sabemos estar presente nos dados. Vemos que essa probabilidade é razoavelmente alta quando o números de eventos no fundo é alto.

4.4 Análise nos dados do Observatório Pierre Auger

A análise que fizemos na seção anterior usa um fundo com distribuição uniforme em todas as direções, isto é, esperamos o mesmo fluxo em todas as direções do céu. Quando passamos de análise em simulações para análise em dados obtidos em experimentos, nosso cenário sofre uma alteração im-



(a) $|C|$

(b) $|c|$

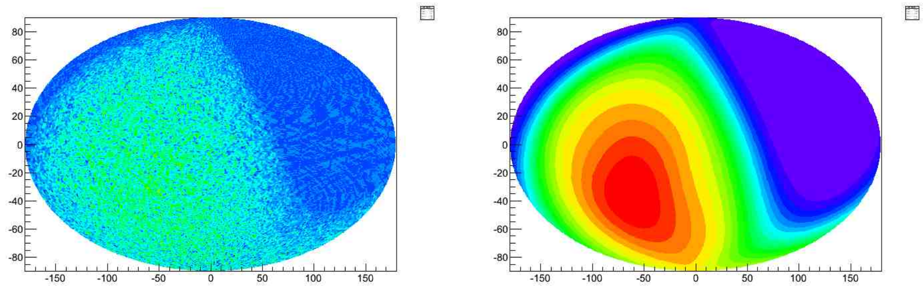
Figura 4.3: (a) Magnitude do maior coeficiente de wavelets para o caso isotrópico (em azul) e isotrópico com um multipleto (em vermelho). (b) Correlação do gráfico $\delta \times 1/E$ para o caso isotrópico e isotrópico com um multipleto.

portante. A exposição dos detectores não é igual em todas as direções e isso significa que essa exposição desigual deve ser levada em conta nas simulações.

Como estamos interessados apenas em valores médios de incidência de raios cósmicos em uma dada direção, podemos ignorar os coeficientes a_{lm} de um sinal para frequências muito altas, que representam apenas detalhes individuais dos raios cósmicos. Uma maneira mais rebuscada de filtrar o sinal é dado pelo uso da transformada inversa de wavelets, como apresentado na seção (3.4). Ignorado as escalas mais baixas na somatória (3.15) eliminamos automaticamente do sinal as estruturas de pequeno tamanho angular. Por exemplo, vemos na tabela 3.1 que usando apenas $j \geq 5$, eliminamos todas as estruturas angulares menores que $11°$. Filtrando o sinal dessa forma também viabiliza um tratamento estatístico mais elaborado em espaço de wavelets.

Na figura 4.4 vemos um exemplo de mapa de cobertura (b) obtido a partir de (a) para partícula com energia $E > 20$ EeV. O mapa de cobertura informa o fluxo de partículas esperado em uma dada direção no céu.

A aplicação do algoritmo apresentado na seção 4.2 nos dados do Observa-



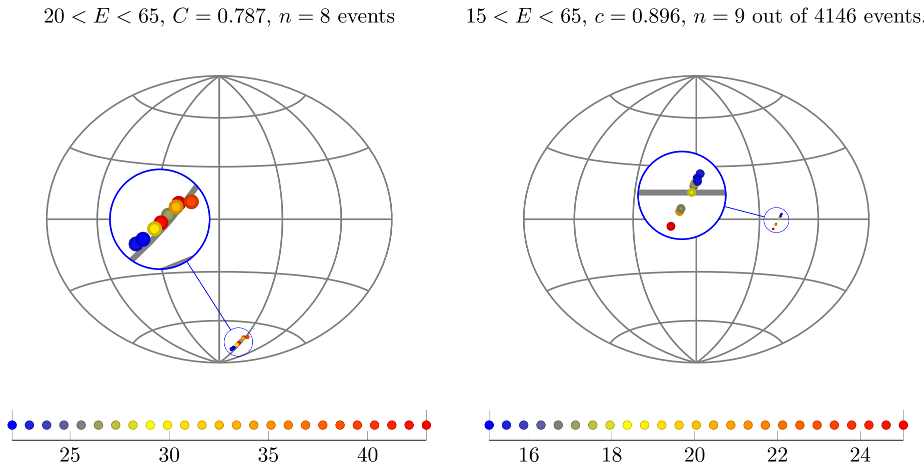
(a) (b)

Figura 4.4: (a) Imagem formada pelos eventos coletados pelo Observatório Pierre Auger para $E > 15$ EeV. (b) Mapa de cobertura formado a partir do mapa (a).

tório Pierre Auger resultou em dois candidatos a multipleto, como apresentado na figura 4.5. As simulações feitas em céus com distribuição isotrópica de eventos, resultou nas seguintes probabilidades de ocorrência ao acaso dos multipletos, $P = 0,47 \times 10^{-3}$ para o multipleto na figura da esquerda e $P < 4.5 \times 10^{-5}$ para o candidato do lado direito da figura.

É difícil comparar esses resultados com os valores obtidos em [10] uma vez que o maoir multipleto encontrado usando a análise de wavelets contém nove partículas. O nosso limite inferior para esse número é de dez eventos. Na opinião do autor da tese, a dificuldade em encontrar candidatos com mais eventos está associada ao fato de que o algoritmo proposto, baseado em wavelets esféricos usa três valores de corte para selecionar os eventos, os coeficiente C que dá a magnitude do coeficiente de wavelet, o coeficiente c que computa a correlação do gráfico $\delta \times 1/E$ e o número de eventos n que compõem o multipleto.

Outra diferença importante do algoritmo [10] está no fato de que em todos os seus passos, sempre é usada uma análise evento a evento, enquanto que a análise de wavelets gera primeiramente uma imagem do céu, formada a partir do número de eventos incidentes em cada pixel. Dessa forma fica evidente



(a) (b)

Figura 4.5: Candidatos a multipletos. Usamos EeV para unidades de energia.

que para que um multipleto seja identificado ele deve necessariamente se destacar em relação ao fundo. Por exemplo, se os pixeis onde incidem os eventos pertencentes ao multipleto contiverem números de eventos parecidos com os outros pixeis da vizinhança, terão também a mesma cor, sendo assim impossível identificá-los.

Capítulo 5

Conclusão

O ESCOPO principal desse trabalho foi de mostrar a utilidade de wavelets esféricos em buscas por estruturas filamentares chamadas multipletos em mapas celestes formados a partir de direções de incidência de raios cósmicos. Esse parece ser o primeiro trabalho que usa wavelets para encontrar multipletos. A grande maioria das aplicações de wavelets esféricos que encontramos estavam relacionadas a estudos de anisotropia em especial em dados de radiação cosmológica de fundo.

Os testes do algoritmo proposto foram feitos em simulações e esses resultados posteriormente foram comparados com dados reais do Observatório Pierre Auger. Os wavelets se mostraram muito sensitivos a multipletos e tiveram um bom desempenho até frações da ordem de 10^{-2} do número de eventos pertencentes ao multipleto com o fundo.

Apêndice A

Funções $d^l_{mn}(\beta)$

VAMOS DAR atenção especial às funções $d^l_{mn}(\beta)$ devido a sua importância na análise de funções definidas em SO(3). Para mais detalhes ver [17]. Essas funções satisfazem as relações de recorrência em três termos

$$\begin{aligned} d^l_{mn}(\beta) &= \frac{2[n-(m+1)\cos\beta]}{[(l-m)(l+m+1)]^{1/2}\sin\beta} d^l_{(m+1)n} \\ &\quad - \left[\frac{(l-m-1)(l+m+2)}{(l-m)(l+m+1)}\right] d^l_{(m+2)n}, \end{aligned} \tag{A.1}$$

com valores iniciais dados por

$$d^l_{0n}(\beta) = (-1)^{l-n}\left[\frac{(2l)!}{(l-n)!(l+n)!}\right]^{(1/2)}\left(\sin\frac{\beta}{2}\right)^{l-n}\left(\cos\frac{\beta}{2}\right)^{l+n}. \tag{A.2}$$

As seguintes relações de simetria também são satisfeitas:

- Simetria com respeito à diagonal ou centro

$$d^l_{-m-n} = d^l_{mn} = (-1)^{m-n} d^l_{mn}. \tag{A.3}$$

- Simetria com respeito à coluna central

$$\begin{aligned} d^l_{mn}(\beta \pm \pi) &= (-1)^{l-n} d^l_{m-n}, \\ &= (-1)^{l-m} d^l_{-mn}. \end{aligned} \tag{A.4}$$

- Para $\beta = \pi/2$

$$\begin{aligned} d^l_{mn}(\pi/2) &= (-1)^{l-m} d^l_{m-n}(\pi/2), \\ &= (-1)^{l-n} d^l_{-mn}(\pi/2). \end{aligned} \tag{A.5}$$

- Paridade

$$d^l_{mn}(-\beta) = (-1)^{m-n} d^l_{m-n}(\beta). \tag{A.6}$$

- Periodicidade

$$d^l_{mn}(\beta + 2\pi) = d^l_{m-n}(\beta). \tag{A.7}$$

A.1 Cálculo de d^l_{mn} via FFT.

O cálculo de $d^l_{mn}(\beta)$ através do algoritmo FFT é a base do algoritmo *Fast Rotational Matching* (veja seção 2.4) e também da nossa implementação da transformada de wavelets na esfera. A forma adotada para o cálculo dessas funções baseia-se na simplicidade e na estabilidade numérica. Basicamente, existem duas formas de cálculo. Uma delas usa relações de recorrência em três passos para calcular o valor de $d^l_{mn}(\beta)$ para um β desejado. A segunda delas usa a transformada rápida de Fourier (FFT), que foi o método adotado nesse trabalho.

A principal vantagem em se usar FFT está em sua estabilidade numérica e performance. Além de FFT continuar um tópico ativo em cálculo numérico, existem muitas implementações de qualidade disponíveis. Se lembrarmos que na tesselação ECP apresentada na seção 2.5, as três coordenadas α, β e γ são igualmente espaçadas, vemos que podemos estender o uso de FFT para todas as variáveis de integração.

A decomposição das funções $d^l_{mn}(\beta)$ em uma série de Fourier pode ser feita notando que uma rotação por um ângulo β pode ser decomposta da forma

$$\mathbf{R}(0,\beta,0) = \mathbf{R}(-\pi/2,0,0)\mathbf{R}(0,-\pi/2,0)\mathbf{R}(\beta,0,0)\mathbf{R}(0,\pi/2,0)\mathbf{R}(\pi/2,0,0). \tag{A.8}$$

Com isso, se denotarmos $\Delta^l_{um} = d^l_{mn}(\pi/2)$, podemos escrever a decomposição de Fourier dos d's da forma

$$d^l_{mn}(\beta) = \sum_{u=-l}^{l} B^l_{mnu} e^{iu\beta}, \tag{A.9}$$

onde

$$B^l_{mnu} = i^{n-m} \Delta^l_{um} \Delta^l_{un}. \tag{A.10}$$

Da definição acima segue imediatamente que

$$B^l_{mn-u} = (-1)^{n-m} B^l_{mnu}. \tag{A.11}$$

Além disso, os coeficientes B^l_{mnu} têm a propriedade

$$B^l_{mn-u} = B^{l*}_{mnu}. \tag{A.12}$$

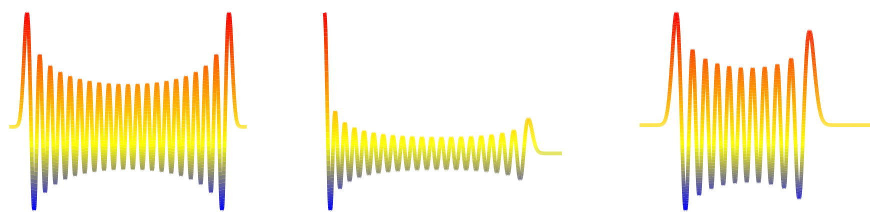
Os coeficientes Δ^l_{mn} satisfazem as equações de recorrência

$$\Delta^l_{l,0} = \left(\frac{2l-1}{2l}\right)^{1/2} \Delta^{l-1}_{l-1,0}, \tag{A.13}$$

$$\Delta^l_{l,n} = \left[\frac{(2l-1)l}{2(l+n)(l+n-1)}\right]^{1/2} \Delta^{l-1}_{l-1,n-1}, \tag{A.14}$$

$$\begin{aligned}\Delta^l_{m,n} \quad = \quad & \frac{2n}{[(l-m)(l+m+1)]^{1/2}}\Delta^l_{m+1,n} - \\ & \left[\frac{(l-m-1)(l+m+2)}{(l-m)(l+m+1)}\right]^{1/2} \Delta^{l-1}_{m+2,n}.\end{aligned} \tag{A.15}$$

Na figura A.1 vemos alguns $d^l mn$



(a) $d^{20}_{0,0}(\beta)$ (b) $d^{20}_{10,0}(\beta)$ (c) $d^{20}_{10,10}(\beta)$

Figura A.1: Funções $d^l_{lm}(\theta)$ de Wigner

A.2 Cálculo explícito para a tesselação usada

Mostraremos alguns detalhes importantes sobre como calcular as funções $d^l_{mn}(\beta)$ usando FFT e mais especificamente, quais os tipos de transformadas devem ser usados no programa FFTW [18].

Para começar é importante lembrar que as funções $d^l_{mn}(\beta)$ satisfazem as relações de simetria (A.3) e portanto são funções ímpares se $m-n$ for ímpar e pares se $m-n$ for par. Com essas propriedades podemos reduzir o cálculo de uma transformada de Fourier para transformadas em senos e cossenos.

As regras de quadratura (2.3), necessárias para a integração de funções

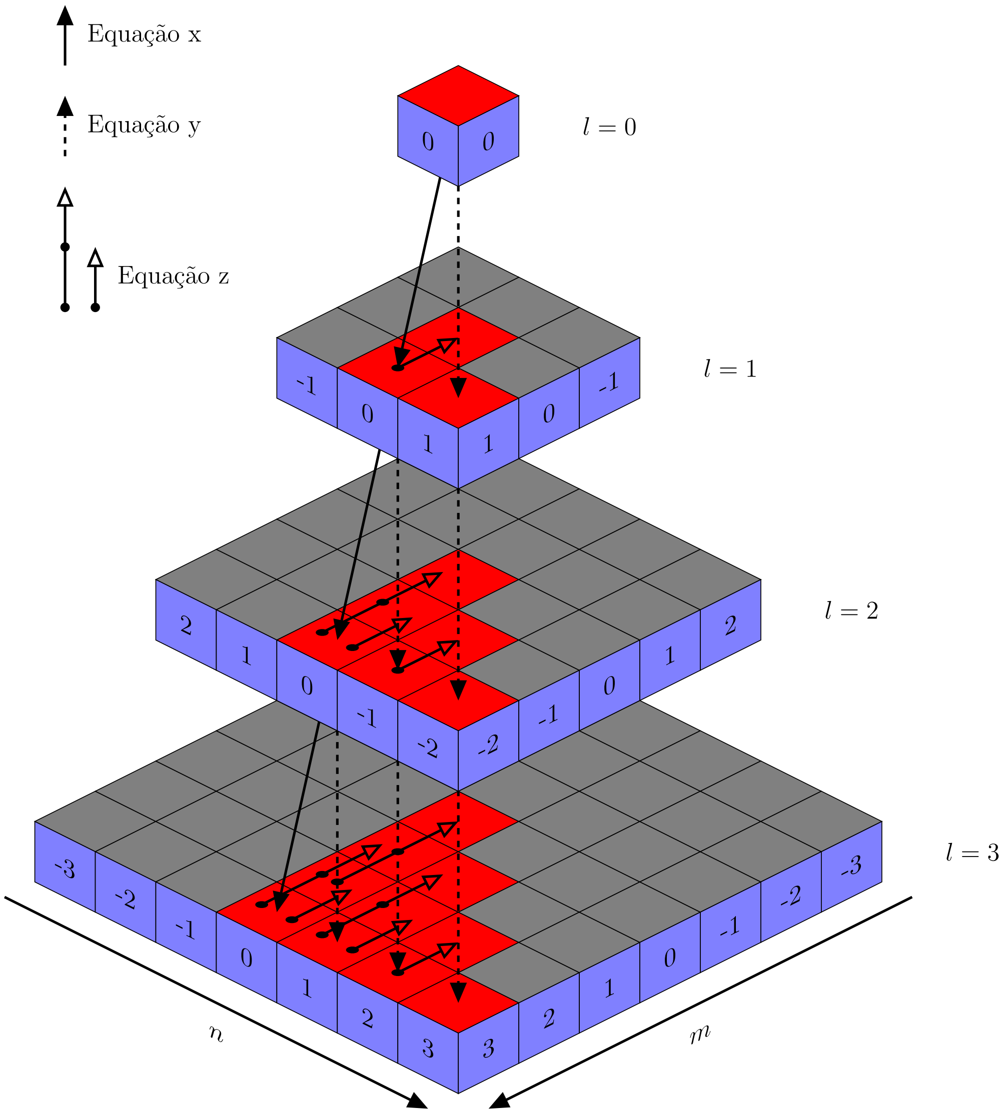


Figura A.2: Cálculo de $d^l_{mn}(\pi/2)$ por recursão
Esquema para o cálculo das relações de recursão. A equação (x) indicada na figura corresponde à equação (A.13), (y) à equação (A.14) e (z) à equação (A.15).

definidas em SO(3), exigem que a coordenada β seja amostrada na malha

$$\beta_k = \frac{(2k+1)\pi}{4B}, \tag{A.16}$$

onde B é o limite de banda do sinal analisado. Com isso podemos trabalhar a equação (A.9) da forma

$$\begin{aligned} d^l_{mn}(\beta_k) &= \sum_{u=-l}^{l} B^l_{mnu} e^{iu\beta_k}, \\ &= \sum_{u=-l}^{-1} B^l_{mnu} e^{iu\beta_k} + B^l_{mn0} + \sum_{u=1}^{l} B^l_{mnu} e^{iu\beta_k}, \\ &= B^l_{mn0} + \sum_{u=1}^{l} (B^l_{mn} e^{iu\beta_k} + B^l_{mn-u} e^{-iu\beta_k}), \\ &= B^l_{mn0} + \sum_{u=1}^{l} (B^l_{mn} e^{iu\beta_k} + B^{l*}_{mnu} e^{-iu\beta_k}), \end{aligned} \tag{A.17}$$

onde B^l_{mnu} está definido em (A.10). Faremos o tratamento separado para os casos em que $m-n$ é par ou ímpar.

Caso $n-m$ par. Partindo da igualdade (A.17) e denotando $n-m=2e$ temos

$$\begin{aligned} d^l_{mn}(\beta_k) &= i^{2e}\left[\Delta^l_{0m}\Delta^l_{0n} + 2\sum_{u=1}^{l} \Delta^l_{um}\Delta^l_{un}\left(\frac{e^{iu\beta_k} + (-1)^{n-m} e^{-iu\beta_k}}{2}\right)\right] \\ &= (-1)^e\left[\Delta^l_{0m}\Delta^l_{0n} + 2\sum_{u=1}^{l} \Delta^l_{um}\Delta^l_{un}\cos\left(\frac{\pi u(k+1/2)}{2B}\right)\right] \end{aligned} \tag{A.18}$$

Na documentação do pacote FFTW vemos que essa transformada tem o tipo **REDFT01** cuja forma geral é

$$Y_k = X_0 + 2\sum_{j=1}^{n-1} X_j \cos[\pi j(k+1/2)/n]. \tag{A.19}$$

Comparando a equação acima com (A.18) vemos que $n = 2B$. Os coeficientes de Fourier são dados por

$$X_j = \begin{cases} (-1)^e \Delta^l_{jm}\Delta^l_{jn}, & j \leq l. \\ 0, & j > l. \end{cases} \tag{A.20}$$

Caso $n-m$ **ímpar.** Partindo da última igualdade em A.17 e denotando $n-m = 2g+1$ temos

$$\begin{aligned} d^l_{mn}(\beta_k) &= i^{2g+1}\left[\Delta^l_{0m}\Delta^l_{0n} + 2\sum_{u=1}^{l}\Delta^l_{um}\Delta^l_{un}\left(\frac{e^{iu\beta_k}+(-1)^{2g+1}e^{-iu\beta_k}}{2}\right)\right], \\ &= 2(-1)^g\sum_{u=1}^{l}\Delta^l_{um}\Delta^l_{un}\sin\left(\frac{\pi u(k+1/2)}{2B}\right), \\ &= 2(-1)^g\sum_{u=0}^{l}\Delta^l_{(u+1)m}\Delta^l_{(u+1)n}\sin\left(\frac{\pi(u+1)(k+1/2)}{2B}\right), \end{aligned} \tag{A.21}$$

onde a segunda igualdade acima é obtida notando que, se usarmos (A.12) e (A.10) podemos mostrar que $B^l_{mn0} = (-1)^{n-m}B^l_{mn0}$, portanto $B^l_{mn0} = 0$ para $n-m$ ímpar.

Na documentação do pacote FFTW vemos que essa transformada tem o tipo **RODFT01** cuja forma geral é

$$Y_k = (-1)^k X_{n-1} + 2\sum_{j=0}^{n-2} X_j \sin[\pi(j+1)(k+1/2)/n], \tag{A.22}$$

Comparando (A.21) com a equação acima obtemos $n = 2B$. Os coeficientes de Fourier são dados por

$$X_j = \begin{cases} (-1)^g \Delta^l_{(j+1)m} \Delta^l_{(j+1)n}, & 0 \leq j \leq l. \\ 0, & j > l. \end{cases} \tag{A.23}$$

Apêndice B

O pacote SWAT

You can't trust code that you did not totally create yourself.
KEN THOMPSON

What I cannot create, I do not understand.
RICHARD P. FAYNMAN

Há muitos anos, computação passou a ser tópico comum aos físicos, sejam eles teóricos ou experimentais. É bastante comum encontrar pesquisadores que passam o dia programando. No caso desse trabalho de doutorado, não foi diferente, a maior parte do tempo do autor foi gasto implementando os assuntos dos capítulos 2, 3 e 4 em linguagem de computador. Dessa forma esse apêndice é dedicado mostrar ao leitor as dificuldades encontradas ao longo do caminho, assim como um manual de como utilizar o programa desenvolvido.

Logo nos primeiros meses do doutorado, percebi que lidar com o problema computacional demandaria muito do meu tempo, principalmente no meu caso, com uma formação até então em física teórica e sem experiência em programação e computação em geral.

Tínhamos o seguinte cenário. Estava bem definido que usaríamos wavelets para localizar multipletos e também qual seria o wavelet, mas tínhamos um problema pela frente. O tempo consumido para executar apenas uma transformada de wavelets era de 72 minutos de acordo com a publicação que introduziu o wavelet que usamos. Isso era preocupante, uma vez que, para validação do método, milhares de simulações devem ser feitas. Se considerarmos por exemplo que em geral teremos que rodar a transformada 10.000 vezes em dados simulados, precisaríamos de 500 dias de simulação. Obviamente poderíamos pensar em dividir o trabalho em várias máquinas, usando um *cluster*, mas isso não melhorava muito as coisas.

Para que fosse possível reduzir o tempo computacional de análise, teríamos que conhecer muito bem os detalhes de implementação da transformada de wavelets. Dessa forma, decidimos ter nossa própria implementação da transformada, que permitiria maior controle e possibilidade de otimização. Assim, começamos a implementação da transformada alguns meses após o início do doutorado. A situação no entanto piorou quando constatamos que a nova implementação, para nosso espanto, levava aproximadamente 5 horas em vez de 72 minutos.

O problema estava no fato de estarmos usando um algoritmo com complexidade $O(n^5)$ quando matematicamente está provado ser possível resolvê-lo em $O(n^4)$. Foi mais ou menos nesse período que encontramos uma publicação, onde é apresentada uma relação de recorrência que permite o cálculo da funções d^l_{mn} de Wigner via FFT.

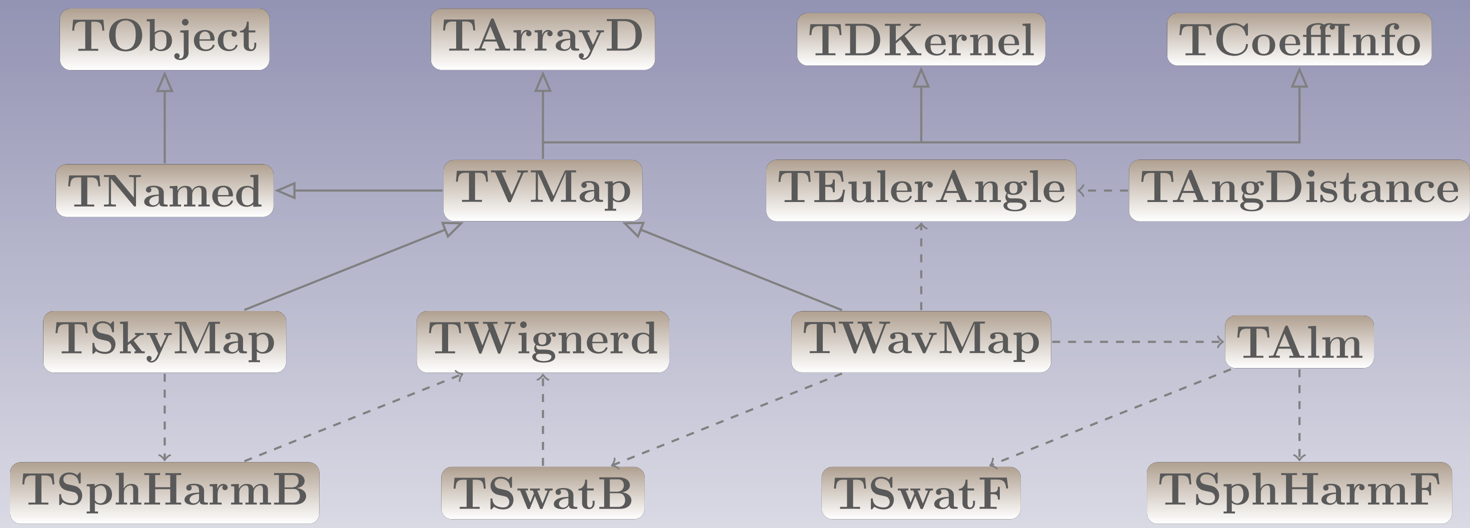
Após implementá-lo com sucesso, o tempo de execução caiu consideravelmente, e passou a ser comparável ao tempo de execução mostrado no artigo original sobre wavelets. Mesmo assim, as simulações ainda não seriam possíveis e portanto muito mais trabalho teria que ser feito. Ao longo dos vários anos seguintes de doutorado, a experiência em computação foi aumentado e o tempo de execução foi se aproximando de algo aceitável. Atualmente temos algo em torno de 3 minutos para a transformada. Todas as simulações apresentadas nessa tese foram feitas usando apenas uma máquina, de forma sequencial, sem precisar usar vários núcleos, que estão comumente disponíveis nas máquinas modernas, ou em *clusters* de computadores, onde o tempo de uso é muito mais caro.

A implementação ainda não se encontra em seu *estado da arte*. Existem muitos pontos que podem ter melhorias significativas. Eu não ficaria surpreso se o tempo total de execução pudesse ser reduzido para apenas alguns segundos. Isso no entanto exigiria uma reestruturação do código para a qual não há tempo nesse doutorado.

É importante lembrar aqui que o desenvolvimento de um programa de computador demanda muitos conhecimentos, que não fazem parte dos cursos de física em geral. Por exemplo, ordenação, busca, árvores, listas, números aleatórios, aritmética de ponto flutuante, arquitetura de computadores, compiladores, *linkers* etc. passaram a fazer parte do meu cotidiano.

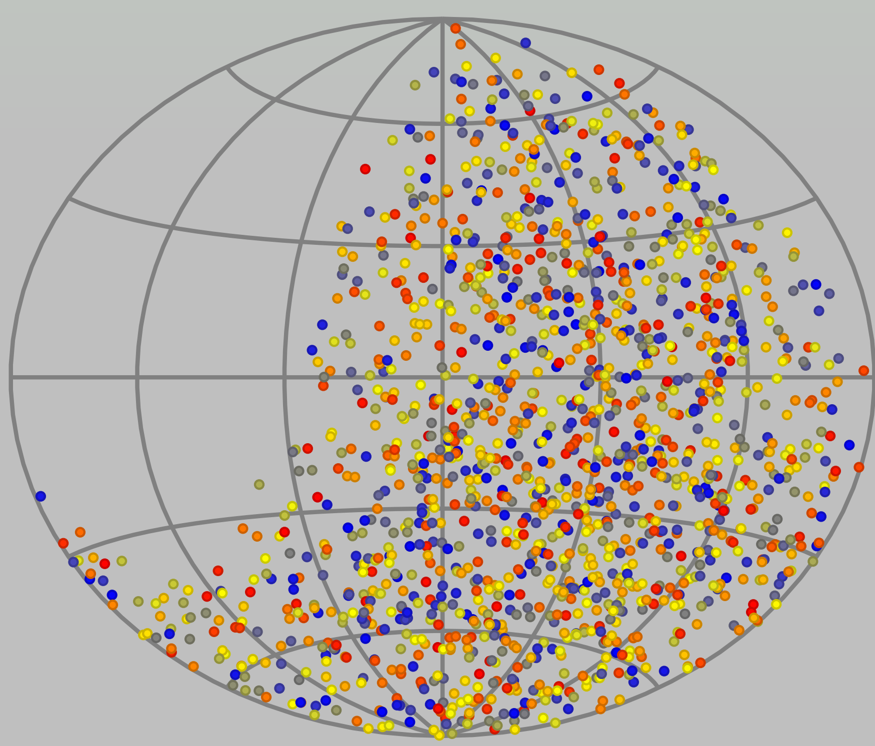
Pode parecer exagero mencionar esses tópicos aqui já que eles parecem “simples"o suficiente para uma pessoa com formação em ciências exatas dominá-los rapidamente, mas quando se trata de um programa que precisa de alto desempenho, é preciso entrar mais a fundo nesses tópicos. Obviamente é possível fazer uma programa que “funcione", mas isso pode vir ao custo de horas e horas a mais em um *cluster*, que não toma tempo apenas do pesquisador em questão, mas de todos os usuários.

O código que desenvolvi para análise dos dados acabou resultando em um pacote, que está disponível sob a licença GPL(GNU Public Licence). Decidimos então tornar esse pacote uma parte do trabalho de doutorado e esperamos que ele possa beneficiar outros pesquisadores que precisem de alguma forma da análise de wavelets em suas pesquisas. O que segue é o guia do usuário do programa SWAT.



SWAT

The Spherical Wavelet Analysis Tool



Manual for version 1.53 - Marcelo Zimbres

Contents

1 Introduction

SWAT is a package for analysis of functions that are defined on the sphere. It has been used extensively to analyze data collected by the Pierre Auger experiment, CRPropa simulations and other simulation data. The program has been written from scratch as part of my my Ph.D, where it evolved from a simple ROOT macro and got bigger and bigger, since the code may be useful to other people I decided to organize, document it and make it public. The project main feature is its C++ implementation of the spherical wavelet transform as presented in [2]. Some attractive features of SWAT are:

- Harmonic and Wavelet transform on the sphere.
- Interface to Healpix code, which is included in the build system (the user does not have to install it).
- Easy selection of events hitting sky window.
- Calculation of deflection vs. $1/E$ graphs.

- ROOT and FFTW are the only prerequisites.

The package includes some parts of the Healpix code. There are two reasons why I decided to include it here, instead of just link against Healpix libraries.

1. I do not need all Healpix routines and support to the fits format.
2. I usually need shared libraries to call Healpix code in a ROOT session, which are not built by Healpix build system since most of its code are C++ templates.
3. Healpix installation used to be messy.

Most of the theoretical details of analysis of functions defined on S^2 and $SO(3)$ were taken from [2] and references therein.
The code has been tested in Linux and MacBook, but the code is portable enough to be built on other platforms. In the following we describe how to use SWAT. Questions concerning the software can be sent to Marcelo Zimbres mzimbres@gmail.com

Acknowledgements

Most of SWAT were developed with financial support from CAPES. Additionaly I would like to acknowledge the Institute of Physics at UNICAMP and the Bergische Universität Wuppertal for finacial support, Brunel University for granting me a scholarship to participate in the Cern School of Computing, the Ettore Majorana Foundation and Center for scientific Culture, the brasilian group of the Pierre Auger experiment, with special thanks to my supervisor Ernesto Kemp, for beliving in the project and Rafael Alves Batista for testing the code. The Pierre Auger Group at university of Wuppertal, with special thanks to Nils Niertenhöfer for helping me with CRPropa.

1.1 Installation

As a prerequisite to install SWAT you need ROOT and FFTW installed, the configure script will fail if they are not installed. SWAT uses autotools to generate the configure script and the Makefile, so you can expect all the standard configure options and makefile targets. The lines bellow will install the libraries on "/usr/local"

```
$ tar -xvzf swat-1.53.tar.gz
$ cd swat-1.53
$ ./configure CXXFLAGS="-O3 -ffast-math"
$ make
$ sudo make install
```

The flags passed to the configure script are optional, but greatly improve performance. To use some of the features of the package, I am assuming you are able to build shared libraries on your platform. To easy the task of loading swat libraries the macro load.C, will be installed on "/usr/local/share/swat". To load the libraries in ROOT's C++ interpreter you have to execute it in your ROOT session, or add it to your code .rootlogon.C macro.

```
$ cp /usr/local/share/swat/load.C ~/.rootlogon.C
$ root # The .so's are automatically loaded here.
```

All SWAT classes are now available in the ROOT session with syntax highlighting (currently I do not use swat from inside a root session, but it may be usefull for others). You should also be able to generate documentation in HTML format with the macro "prefix/share/swat/makehtml.C". You have to run this macro from the directory where you built SWAT. The documentation will be built in the directory htmldoc, this is a nice way to get aquainted with the code (this is meant for those wanting to develop).

1.2 Getting ready

To use Herald file (a file distributed by the Pierre Auger collaboration containing collected data), you will have to convert the ascii file to a **TTree** and save it in a .root file, for that you should use the macro prefix/share/swat/-convert_herald.C where prefix is usually /usr/local. Copy this macro to the same directory where you have the herald data file. The macro will read a file with name "herald.dat", so you may have to rename your file or edit the macro. For CRPropa simulations you can configure CRPropa to output a .root file instead of a text file. The code uses the title of the **TTree** to differentiate between a Herald and CRPropa file. The title of the TTree for CRPropa must be "CRPropa 3D events", as far as I know, this is the default.

1.3 High quality graphs with pgfplots

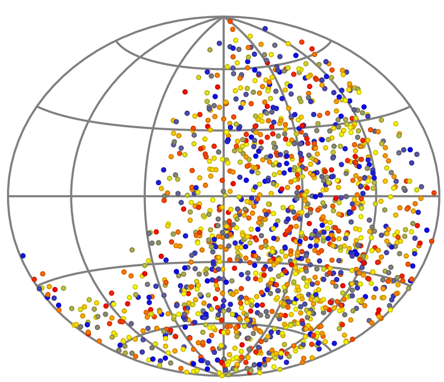
If you have pgfplots installed on your machine, you can use the LaTeX files which are available in the directory pgfplots in swat root dyrectory, to generate high quality graphs using some output of swat analysis. You may have noted that after running swat_find you get many text files in your working directory. They are input files for the LaTeX files. In the following subsections we show how to use the material.

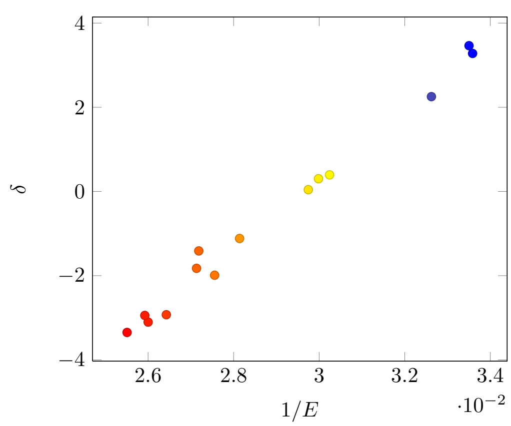
1.3.1 Skymaps

When you run the program swat_find or swat_sim, they produce a file named *skymap.dat*. This file contains the coordinates and energy of the events that passed the cut. To generate a graph for an isotropic sky for example, one can use:

```
$ tar -xzf skymap.tar.gz
$ cd skymap
$ swatsim -n 1000 -s 1
$ pdflatex --jobname=skymap-f1 skymap.tex
```

Example skies can be seem here:





The program will also generate the file mult_cand.tex, this file contains only the events which hit the tangent plane in the position found by the wavelet analysis. Usually the energy of events is represented in skymaps as circles of varying size, where events with higher energy have larger circles. I do not like such representations since they can give the false impression that the angular resolution gets worse as the energy increases, which is not true. Additionally, these circles pollute the figure. In the representations I am using, colors are used to differentiate energies which in my opinion is a much clear way of representing the sky.

1.3.2 Energy deflection graphs

In addition to skymap.dat, the files corr_graphN.dat will be also generated, with N varying from 0 to 15 (this number can be passed in the command line). These files contains the energy-deflection graphs, of the events the hit the tangent plane in the position found by the wavelet analysis. On figures above, we see one example. The graph on the left originates from a CRPropa simulation.

1.3.3 Wigner-d functions

Swat implements the calculation of the wigner d^l_{mn} functions via FFT. You can use the file pgfplots/wignerd.tar.gz to generate some nice graphs.

2 Transforms on the Sphere

Most SWAT functionality is based on Fourier and wavelet transforms on the sphere, therefore we give in this section a brief overview of analysis of functions on the sphere. Before that however, it is important to mention the problem of sampling functions on the sphere, which is refered to as its pixelization. Due to the fact that most of the time we work with function defined on the line (temporal series, for example) or on the plane (images in general) we do not have to care much about how to sample. We simply divide it in squares of same area and everything works just fine. However, when we try the same thing on the sphere we imediately see that it is impossible to achieve same area pixelization on the sphere due its curvature.
There are two main features we would like to have when dividing a sphere in pixels. First we would like to have pixels of same area, this is achieved by the Healpix pixelization 2. Second, we would like to to use FFT to calculate functions on the sphere instead of slow recursion and that is achieved by ECP Pixalization, which stands for "Equidistant cylindrical projection"2. Unfortunately theses two features can not be achieved at the same time.

In this software we decided to use ECP due to its simplicity. In ECP, the three Euler angles, which we are using to parametrize SO(3), are sampled as follows α, β and γ

$$\alpha_{j_1} = \gamma_{j_1} = \frac{2\pi j}{2B}, \quad \beta_{j_1} = \frac{\pi(2k+1)}{4B} \tag{1}$$

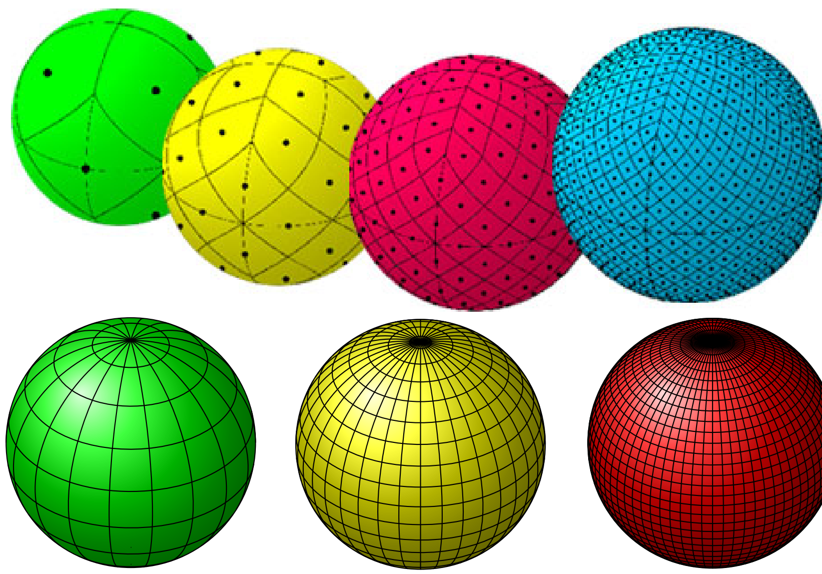


Figure 1: Above:Healpix pixelization of the sphere. Below: The ECP pixelization.

2.1 Fourier Transform on SO(3)

A function $f \in L^2(SO(3))$ can be decomposed as follows

$$f(\alpha,\beta,\gamma) = \sum_{l=0} \sum_{m=-l}^{l} \sum_{n=-l}^{l} f^l_{mn} D^l_{mn}(\alpha,\beta,\gamma). \tag{2}$$

where

$$D^l_{mn}(\alpha,\beta,\gamma) = e^{-im\alpha} d^l_{mn}(\beta) e^{-in\gamma}. \tag{3}$$

is called "Wigner-D functions" and form an irreducible representation on SO(3). The functions d^l_{mn} are called "small wigner-d functions".

The inverse of 2 on ECP pixelization is given by

$$f^l_{mn} = \frac{1}{(2B)^2} \sum_{j_1=0}^{2B-1} \sum_{j_2=0}^{2B-1} \sum_{k=0}^{2B-1} w_B(k) f(\alpha_{j_1}, \beta_k, \gamma_{j_1}) D^{l*}_{mn}(\alpha_{j_1}, \beta_k, \gamma_{j_1}) \tag{4}$$

where B is the band limit and the quadrature weights are given by

$$w_B(k) = \frac{2}{B} \sin\left(\frac{\pi(2k+1)}{4B}\right) \sum_{j=0}^{B-1} \frac{1}{2k+1} \sin\left((2j+1)(2k+1)\frac{\pi}{4B}\right) \tag{5}$$

where $0 \leq k < 2B$.

3 Installed programms

In the following sub-sections we will describe the main programs which SWAT installs and which provide an easy interface to the wavelet analysis.

3.1 swat_find

swat_find is the main program. It can be used to find sources in Pierre Auger data and CRPropa simulations. Beyond that, it can be also used to test whether the events belong to a multiplet, calculating energy-deflections graphs for the sources found. All information is saved in .root format. A source here is a term used to mean something with a position on the sky, usually described by (θ, ϕ) and an orientation. We will use the three Euler angle (α, β, γ) to describe the source.
It uses the following algorithm:

1. The TTree containing Pierre Auger data or CRPropa simulations is read from the file passed in the command line with option -f.

2. Events with energy in the range $emin < E < emax$ are selected and used to fill a Healpix map. The options -i and -e are used to pass the energy range.

3. The Healpix map generated is transformed to wavelet space. The parameter N, passed by the option -N gives the precision on the ability of the wavelet to find the angle γ, which can range from 1 - 128 in our implementation. The precision in degrees are calculated with the

formula $180/N$ The scale on which the analysis is performed is passed with option -j. It ranges from 0 - 8. We found out that the best results are achieved with j = 1.

4. A partial sort is used to find the 15 largest wavelet coefficients.

5. The euler angles found will be used to calculated the tangent plane equation for each source. The width and length of the plane are passed with options -w and -l respectively.

6. A loop on the data is made to select all events hitting the tangent plane. The energy cut is used again.

7. The energy-deflection graphs are calculated and the correlations found are printed on the screen.

All the information is saved in the file "sources.root". Lets see an example:

```
$ swatfind -j 1 -N 127 -i 20 -e 40 -w 2 -l 10 -f chain.root
TFile**         chain.root      CRPropa output data file
 TFile*         chain.root      CRPropa output data file
  OBJ: TNtuple  events  CRPropa 3D events
  OBJ: TEventList       list    20 < emin && emax < 40
  OBJ: THealpixMap      hmap    Healpix sky map
  OBJ: TGraph   g0      C = -0.993, N = 4
  OBJ: TGraph   g1      C = -0.986, N = 5
  OBJ: TGraph   g2      C = 0.000, N = 0
  OBJ: TGraph   g3      C = -1.000, N = 2
   ...
```

The N in the TGraph title is the number of events in the graph (or the number of events which hit the tangent plane) and is not to be confused with the N passed by command line option -N. C is the correlation of the energy-deflection graphs.
Now, if you want to see the exact location of the source:

```
$ root sources.root
root[1] .ls
TFile**         sources.root
 TFile*         sources.root
  KEY: THealpixMap      hmap;1  Healpix sky map
```

```
  KEY: TEulerAngle      source0;1
  KEY: TEulerAngle      source1;1
  KEY: TEulerAngle      source2;1
root [2] source0->Show(1)
wav(137.812,-27.4219,159.638) = -0.035312
root [3]
```

Help description

```
Searches for multiplets in herald data or CRPropa
simulations. Calculates correlation of events hitting
stripe on the tangent plane and dispalys on the screen.
The results are saved to sources.root file. Two input
file formats are supported, both are TTrees saved in a
.root file. The TTrees can be either the output of
CRPropa or a Herald file converted to TTree (see macro
macros/convert_herald.C in swat source tree).

Usage: swat_find [ -j scale] [-N number] [-i emin]
       [-e emax] [-w width] [-l length] [-f file.root]
       [-n nsources] [-t wav_threshold]

Options:

-h:     This menu.
-j:     Wavelet scale. It is a number in the range
        0 <= j <= 8, defaults to 1.
-N:     Band limit of wavelet, in the range 0 < N <= 128,
        defaults to 1.
-i:     Minimum energy of events, defaults to 20 EeV.
-e:     Maximum energy of events, defaults to 40 EeV.
-w:     Width of tangent plane, defaults to 2 degrees.
-l:     Length of tangent plane, defaults to 10 degrees.
-f:     Root file containing Tree with data, defaults to
        chain.root.
-n:     Number of sources to look for. Default to 15
-t:     Wavelet threshold value.
```

3.2 swat_sim

swat_sim uses Monte Carlo simulations to calculate the probability of a multiplet happen by chance on a isotropic sky. The distribution of E, θ and ϕ

must be passed in the command line, it can be generated by swat_gen The probability is calculated as follows:

- The program simulates n isotropic skies. If you want to add your simulated events to each sky you can use the option -f and a TTree in CRPropa format will be read and added in each sky.
- For each sky simulated the analysis made by swat_find is used to calculate the correlation coefficient.
- If the correlation is larger than C , passed with option -c and has more than n events, passed with option -m, then the multiplet has passed the criteria.
- The probability will be total number of multiplets passing the criteria divided by number of skies simulated.

Beyond the probability, the program outputs two additional histograms:

1. A histogram of the correlation coefficients found, for which the total number of events is larger than the value passed with -m.
2. A histogram of the number of events in the tangent plane.
3. A histogram of the largest wavelet coefficients found.

Help message

```
Calculates the probability of a multiplet with minimum
correlation c > c_0 (see -c option), minimum number of
events m > m_0 (see -m option) and where the magnitude of
the wavelet coefficint e C > C_0 (see -C option), happen
by chance using wavelet analysis.  First an isotropic sky
is simulated (the coverage and energy distribution must
be provided) and the wavelet representation of the sky is
calculated, the euler angles of the largest coefficient
is used to calculate the equations of the tangent plane
at the position found (the euler angles). The correlation
c is calculated including all events that hit the tangent
plane, whose size is specified with the options -l and
-w. The probablility will be the number of multiplets
with c > c _0, C > C_0 and m > m_0, divided by the number
```

```
of skies simulated. Additionaly, seven other quantities
are calculated:

   1 - The histogram of the number of events that hit the
       tangent plane.
   2 - The histogram of the c's found for which the
       number of events is greater than m_0 and C > C_0
       (passed in the command line).
   3 - The histogram of the magnitude of wavelet
       coefficients C_0.
   4 - The histogram of the mean of wavelet coefficients.
   5 - The histogram of the variance of wavelet
       coefficients.
   6 - The histogram of the skewness of wavelet
       coefficients.
   7 - The histogram of the kurtosis of wavelet
       coefficients.

If -f option is used, a TTree in the file will be read
and events will be added to the analysis, this is useful
to include a simulated multiplet on the analysis, hiding
it in the isotropic backgroung the test the algorithm.

It is also mandatory to specify:

   - Energy distribution.
   - The theta distribution.
   - The phi distribution.

This is the distribution the background events have to
follow. These distributions are read from a root file.
Use swat_gen and swat_coverage to generate them.

Usage: swat_sim [ -j scale] [-N number] [-n nevents]
       [-s skies] [-i emin] [-e emax] [-c corr] [-m mevents]
       [-w width] [-l length] [-f file.root] [-C min_wav]
       [-d energy] [-a coverage]

Options:

-h:     This menu.
-j:     Wavelet scale, a number in the range 0 <= j <= 8,
        defaults to 1.
-N:     Band limit of wavelet, in the range 0 < N <= 128,
        defaults to 1.
```

```
-n:       Number of events in the simulated sky, defaults
          to n = 1000
-s:       Number of skies to simulate, defaults to 100.
-i:       Minimum energy of events, defaults to 20 EeV.
-e:       Maximum energy of events, defaults to 40 EeV.
-c:       Minimum correlation, defaults to 0.2.
-C:       Minimum Magniftude of wavelet coefficient,
          defaults to 0.0.
-m:       Minimum number of events hitting tangent plane.
-w:       Width of tangent plane, defaults to 2 degrees.
-l:       Length of tangent plane, defaults to 10 degrees.
-f:       Add events in TTree stored in file to the
          simulated sky.
-d:       Histogram with energy distributions.
-a:       Histogram with theta and phi distributions.
```

3.3 swat_gen

Generates distribution of E θ and ϕ coordinates that will be used by swat_sim program. Two kinds of distribution are supported, isotropic or Auger distribution if herald data is provided.

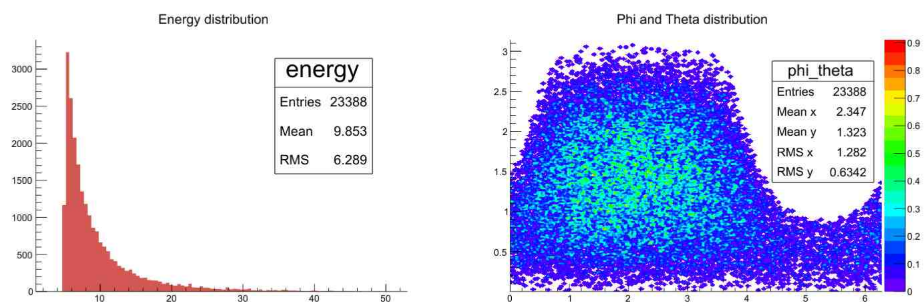


Figure 2: Energy distribution on the left. On the right angular distribution. Both for Auger experiment.

3.4 swat_prob

```
Reads a file output by swat_sim and calculates
probabilities.  Use the macros cprob.tex and
```

```
wprob.tex in pgfplots directory to produce
probability graphs.

Usage: swat_prob [-n n_steps] [-o outpu-file] [-t]
                 [-k kind]

Options:

  -h:      This menu.
  -o:      Output file , where probabilities
           will be recorded.
  -t:      Calculates P(q >= q_0) if provided,
          P(q < q_0) otherwise.  q will be the
           correlation or wavelet coefficient,
           depending on the value passed in option -k
  -k:      Either 1 for correlation or two for
           wavelet coefficient.
  -n:      Total number of steps (points on
           probability graph). Defaults to 10.
```

3.5 swat

This program is only used to benchmark and test the algorithm. It can test both the spherical harmonic transform and the spherical wavelet transform.

```
$ time swat -J 8 # Will test spherical harmonic transform.
$ time swat -J 8 -N 3 # Will test wavelet transform.
```

Help message

```
Tests the algorithm performing forward and backward
transform.  Both spherical harmonic transform (if option
-N is not provided) or spherical wavelet transforms can
be performed. I use this program to benchmark my code
using the time command:

$ time swat -J 8 -N 127

for example. If forward folloed by backward transform do
not result in the data, with precision 1e-10, program
exits with EXIT_FAILURE status. For example
```

```
$ swat J8 -N 127
$ echo $?
0
$

Usage: swat [-J j] [-N n]

Options:
-h:      This menu.
-J:      Sets band limit of the signal to 2^J, defaults to
         J = 7.
-N:      Band limit of wavelet to be used.
```

References

[1] J Fourier Anal Appl (2008) 14: 145179.

[2] Mon. Not. R. Astron. Soc. 000, 122 (2007).

[3] http://www.ifi.unicamp.br/~mzimbres/

Apêndice C

Artigos

Alguns resultados originais dessa tese foram publicados em periódicos científicos, um deles nacional e outro internacional, ambos em inglês. O primeiro deles *Identifying Patterns on Cosmic Ray Maps with Wavelets on the Sphere* foi publicado na revista da associação dos Pós-Graduandos do Instituto de Física Gleb Wataghin, *Physicae*. O segundo artigo *Using spherical wavelets to search for magnetically-induced alignment in the arrival directions of ultra-high energy cosmic rays* foi publicado no periódico *Astroparticle Physics*.

Using spherical wavelets to search for magnetically-induced alignment in the arrival directions of ultra-high energy cosmic rays

M. Zimbres[a], R. Alves Batista[b], E. Kemp[a]

[a] *Instituto de Física "Gleb Wataghin" - Universidade Estadual de Campinas, 13083-859, Campinas-SP, Brazil*
[b] *II. Institut für Theoretische Physik - Universität Hamburg, Luruper Chaussee 149, D-22761, Hamburg, Germany*

Abstract

Due to the action of the intervening cosmic magnetic fields, ultra-high energy cosmic rays (UHECRs) can be deflected in such a way to create clustered energy-ordered filamentary structures in the arrival direction of these particles, the so-called multiplets. In this work we propose a new method based on the spherical wavelet transform to identify multiplets in sky maps containing arrival directions of UHECRs. The method is illustrated in scenarios with a multiplet embedded in isotropic backgrounds with different number of events. The efficiency of the algorithm is assessed through the calculation of Type I and II errors.

Keywords: spherical wavelets, ultra-high energy cosmic rays, cosmic magnetic fields, multiplets

1. Introduction

Cosmic rays were discovered more than one century ago. One remarkable feature of the cosmic ray spectrum is that it spans more than ten orders of magnitude, up to hundreds of EeV (1 EeV $= 10^{18}$ eV). The spectrum roughly follows an inverse power law, which means that the expected flux of particles at the highest energies is extremely low compared to the lower energies. In fact, at energies of a few EeV, only one particle per square kilometer per year is expected. Particles with energies $\gtrsim 1$ EeV are referred to as Ultra-High Energy Cosmic Rays (UHECRs). Some experiments, such as the Pierre Auger Observatory and the Telescope Array, have been designed to increase the statistics of events in this energy range. Despite the higher statistics achieved, which has shed light over many important questions, some of them still remain unanswered, such as the origin, nature and mechanisms of acceleration of these particles.

Due to the presence of galactic and extragalactic magnetic fields, charged cosmic rays are expected to be deflected. Hence, incoming directions, as measured by a detector, do not point back to the exact position of the source. The magnitude of the deflection depends on the strength of the intervening fields. In the case of charged particles the deflections are roughly inversely proportional to the particle energy. Therefore, for coherent fields, the different Larmor radii described by cosmic rays can create filamentary structures ordered by energy, the multiplets. This allows the reconstruction of the source position, and consequently enhances the possibility to do astronomy with UHECRs.

In this paper we propose a new method of identifying multiplets, based on the spherical wavelet transform. The paper is organized as follows: in section 2, we review the physics underlying multiplets; in section 3 we present the wavelet transform from a pattern matching algorithm point of view, and give motivations for its use in cosmic ray physics; in section 4 we present a novel algorithm to identify filamentary structures in cosmic ray maps; in section 5 the method is applied to simulated data sets; in section 6, we present our results; in section 7 we make our final remarks.

2. Cosmic Magnetic Fields and Multiplets

Several results show that at ultra-high energies the cosmic ray spectrum might contain a component of atomic nuclei. Data from the High Resolutions Fly's Eye Experiment (HiRes) indicate that UHECRs are probably protons [1], whereas the results from the Pierre Auger Observatory indicate that the composition tends to heavy nuclei at the highest energies[1] [2]. Despite this controversy, one can consider that UHECRs are predominantly charged particles and, as so, can be deflected by magnetic fields.

The deflection expected for a UHECR of charge Z due to the regular component of the galactic magnetic field (GMF) is given approximately by [3]:

$$\delta \approx 53^\circ \frac{Z}{E} \left| \int_0^L \frac{d\vec{r}}{\text{kpc}} \times \frac{\vec{B}}{\mu\text{G}} \right| \text{EeV}, \tag{1}$$

where E is the energy of the particle, $\vec{B}$ the magnetic field, and L the travel distance of the particle. Since $|\vec{B}| \sim \mu$G and is coherent over lengths of ~ 10 kpc, typical deflections are $\sim 10^\circ$ for protons of 10 EeV.

Email addresses: mzimbres@ifi.unicamp.br (M. Zimbres), rafael.alves.batista@desy.de (R. Alves Batista), kemp@ifi.unicamp.br (E. Kemp)

[1] Notice that this results strongly depends upon the hadronic interaction model taken into account.

The expected deflection due to the turbulent component is [4]:

$$\delta_{turb} \approx 10^{\circ} \frac{Z \text{ EeV}}{E} \frac{B_{rms}}{\mu G} \sqrt{\frac{L}{\text{kpc}}} \sqrt{\frac{L_c}{50 \text{ pc}}}, \tag{2}$$

where L_c is the coherence length of the field and B_{rms} is the root mean square intensity of the magnetic field.

According to equation (1) the deflection is inversely proportional to the energy of the particle. Therefore it is possible that energy ordered filamentary structures, the so-called multiplets, can be detected in cosmic ray maps as shown in figure 1. Another method to detect filamentary structures in the arrival directions distributions of UHECRs was proposed by Harari *et al.*[5].

An interesting property of the multiplets is that the position of the source can be reconstructed, allowing one to identify UHECRs sources. A method to reconstruct the source position of a multiplet was presented by Golup *et al.* [6] and was used by the Pierre Auger Collaboration to estimate the position of the sources for some possible multiplet candidates [4]. In the aforementioned work, no evidence for the presence of multiplets was found for energies above 20 EeV.

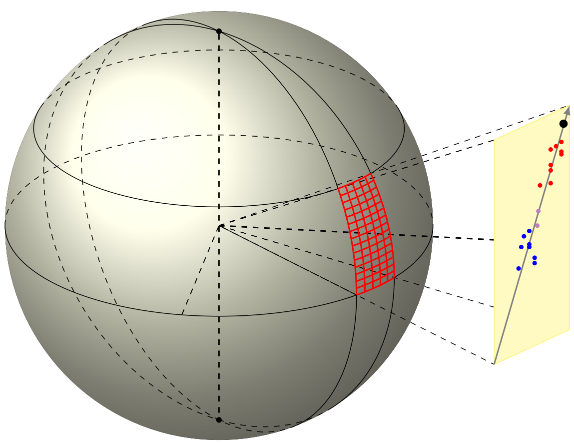


Figure 1: Illustration of a multiplet on the tangent plane of a sky map. The dots represent events from higher (red) to lower (blue) energies. The black dot corresponds to the position of the source.

It is important to notice that for some models of the galactic magnetic fields such as the ones proposed by [7–9] cosmic ray multiplets can be formed, whereas in other models such as the one recently proposed by [10] they are very unlikely to occur, due to the strength of the field, especially the turbulent component.

The role played by extragalactic magnetic fields in the deflection of UHECRs is not fully understood. Simulations of the propagation of UHE particles in the large scale structure of the universe have been performed by several groups [11–14]. However, these results are contradictory and one cannot obtain a clear picture of the effects of the extragalactic magnetic field for the deflection of UHECRs. For the purposes of this work we assume that the deflections induced by extragalactic magnetic fields are much smaller than the ones induced by the GMF.

3. Wavelets on the sphere

In many branches of Physics, specially Astrophysics and Cosmology, wavelets have been successfully applied to solve various problems, particularly related to detection of signals. Wavelets on the plane have been widely used to denoise cosmic microwave background (CMB) maps [15–17]. However, the problem of identifying anisotropies in the distribution of arrival directions of UHECRs has not been properly addressed, and only a few works [18–22] on this topic are available in the literature.

Wavelets are commonly used in one and two dimensional data analysis, but in recent years the interest in data lying on the sphere has increased due to experiments such as the Cosmic Background Explorer (CoBE), the Wilkinson Microwave Anisotropy Probe (WMAP), the Planck Satellite and the Pierre Auger Observatory, which make use of these kind of data. The need to process these sorts of data sets pushed the interest on new techniques under development, being wavelet analysis on the sphere one example. Wavelets on the sphere have been successfully applied to studies of the cosmic microwave background [23–26], and also dark energy detection [27, 28].

Wavelets can be specially useful for cosmic ray data analysis, where we usually have to deal with a non uniform exposure to search for local structures in the sky, with a defined position, such as point sources, and possibly an orientation, such as multiplets. An event by event analysis is not viable, and the analysis in harmonic space would be even harder since all local properties are lost. Spherical wavelets come up as a good alternative to address these problems. Other attractive features of wavelet analysis are:

- any function can be exactly represented by its wavelet coefficients;
- local features of the signal can be enhanced in wavelet domain, meaning that the number of coefficients needed to represent a given signal is reduced;
- it provides scale decomposition, making it possible to identify structures with different angular sizes and focus on the part of interest in the signal;
- it does not rely on any tangent plane approximation, in the case of wavelets on the sphere.

Data representation in wavelet domain can be thought as something between pixel and harmonic representation. Sometimes it is very convenient to decompose the data to enhance properties that are not clear in harmonic or pixel domain. Wavelets can be interpreted as local analysis functions which can be rotated and/or dilated, to obtain information regarding the signal morphology.

3.1. Pattern matching on the sphere

In this subsection we show that the problem of finding a multiplet, or any other pattern defined on the sphere, can be treated by the fast rotational matching algorithm [29].

Let $f(\theta, \varphi)$ and $h(\theta, \varphi)$ be two functions defined on the sphere. Assume that h is a rotated version of f, such that $f = \Lambda(\alpha, \beta, \gamma)h$, with (α, β, γ) being Euler angles, and Λ denoting the rotation operator in $SO(3)$.

If we know that a rotated version of the pattern f is present in h, we can find its latitude, longitude and orientation on the sphere by correlating all rotated versions of f with h, and selecting the rotation which maximizes the correlation

$$C = \int_{\mathbb{S}^2} h(\theta, \varphi)\overline{\Lambda f(\theta, \varphi)}d(\cos\theta)d\varphi. \tag{3}$$

We can parametrize the rotations in terms of Euler angles. In this case the correlation function can be denoted by $C = C(\alpha, \beta, \gamma)$. In other words, we want to find the angles α, β and γ for which C is maximum.

The straightforward evaluation of C is very time-consuming and not affordable depending on the precision we are using to describe our signal. Denoting the band-limit of the signal by B, the complexity of equation (3) is $O(B^5)$. This complexity can be reduced to $O(B^4)$ by calculating the correlation in the harmonic domain. So, we can write the spherical harmonic expansions of f and h as

$$f(\theta, \varphi) = \sum_{l=0}^{B-1} \sum_{m=-l}^{l} a_{lm} Y_m^l(\theta, \varphi) \tag{4}$$

$$h(\theta, \varphi) = \sum_{l=0}^{B-1} \sum_{m=-l}^{l} b_{lm} Y_m^l(\theta, \varphi), \tag{5}$$

with Y_m^l being spherical harmonics. Using these expansions we can show (see Ref. [30] for more details) that the correlation function (3) can be cast as

$$C = \sum_{l=0}^{B-1} \sum_{m,n=-l}^{l} \overline{a_{lm}} b_{ln} D_{mn}^l(\alpha, \beta, \gamma) \tag{6}$$

$$= \sum_{l=0}^{B-1} \sum_{m,n=-l}^{l} \overline{a_{lm}} b_{ln} e^{im\alpha} d_{mn}^l(\beta) e^{in\gamma}, \tag{7}$$

where D_{mn}^l denotes the Wigner-D functions and d_{mn}^l the small Wigner-d functions. This algorithm is called *Fast Rotational Matching*, abbreviated to FRM. It uses the Fast Fourier Transform (FFT) on both α and γ to calculate equation 7. Depending on the tessellation chosen, the FFT can be extended to the β coordinate [29].

This approach to find a multiplet may not be an option because we do not know exactly the pattern $f(\theta, \varphi)$ that describes the multiplet. Moreover, for angular resolutions corresponding to a band limit higher than $B = 128$, too much memory would be required. For $B = 256$, for example, approximately 4.5 GB of RAM would be necessary in our implementation.

In the next subsection we show that the difficulties mentioned above can be overcome by using a special family of directional wavelets instead of the pattern itself.

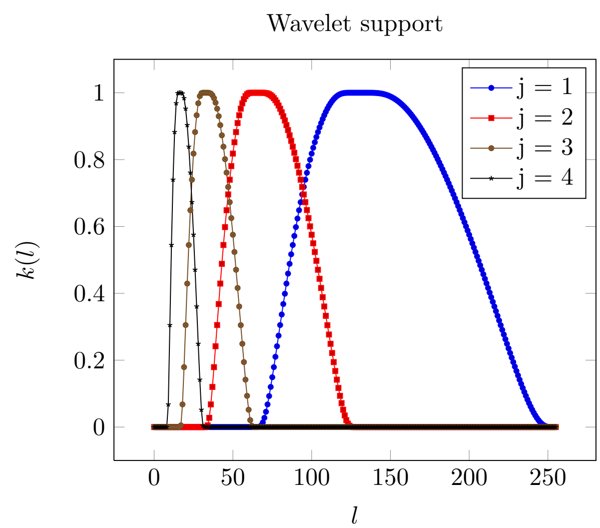


Figure 2: The figure shows the harmonic support in the frequency domain of the family of wavelets used, for $J = 8$.

Table 1: First column: scale j. Second column: wavelet support. Third column: angular sizes in degrees for which the wavelet is sensitive. Fourth column: maximum precision on the angular variables in degrees.

j	Support	Angular size (°)	Precision (°)
0	$(256, 128)$	$(0.7, 1.4)$	0.7
1	$(256, 64)$	$(0.7, 2.8)$	0.7
2	$(128, 32)$	$(1.4, 5.6)$	1.4
3	$(64, 16)$	$(2.8, 11.3)$	2.8
4	$(32, 8)$	$(5.6, 22.5)$	5.6
5	$(16, 4)$	$(11.3, 45.0)$	11.3
6	$(8, 2)$	$(22.5, 90.0)$	22.5
7	$(4, 1)$	$(45.0, 180.0)$	45.0
8	$(2, 1)$	$(90.0, 180.0)$	90.0

3.2. Pattern matching with directional wavelets

When the function f from equation 4 is a wavelet, and h is the signal of interest, equation 3 is the definition of the forward spherical wavelet transform, and the correlation function can be interpreted as the wavelet representation of the signal h.

The wavelets used in this work are defined in the harmonic domain[2]. Their harmonic representation has the special property of allowing them to be split into a kernel and a directional part

$$b_{lm} = k(l) S_{lm}, \tag{8}$$

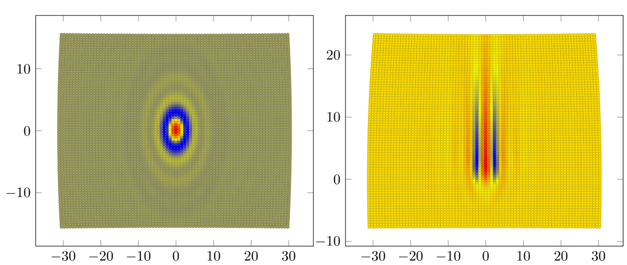
where the kernel $k(l)$ is responsible for dilations and S_{lm} is responsible for the directional properties of the wavelets. This split ensures that dilations do not affect directional properties, so that these two parts can be treated independently.

[2] For the explicit definition and derivation of the mother-wavelet equation, refer to Ref. [31].

The kernel $k(l)$ at each scale j has zero values at frequencies outside the range $(2^{J-1-j}, 2^{J+1-j})$, which is usually referred to as the wavelet support. The support is related to the frequencies to which the wavelets are sensitive. For a graphical representation see figure 2. It is interesting to notice in this figure that an adequate choice of j can suppress some values of l, making this method powerful even when we consider the exposure of a detector, which is usually associated to low values of l.

In our implementation, for a given band limit B, the number of scales in the wavelet analysis is given by $J = \log_2 B$. Since we have used $B = 256$, then $J = 8$. For a physical interpretation, however, it is simpler to think in terms of angular sizes. For that, we can convert the frequency range into an angular size using the formula $360°/2l$, where l is the frequency we are interested in. In table 1 the angular size to which each scale is sensitive is shown.

The sensitivity of the wavelet on finding the angle γ can be controlled by imposing a band limit on S_{lm}. Denoting this band limit by N, then $S_{lm} = 0$ for all $m \geq N$. If the directional features of the signal are not relevant for the analysis, low values of N, such as $N = 1$, can be used. The maximum value of N at scale j is given by $N = 2^{J-j+1}$. The precision of the orientation for a given N is given by $\Delta\gamma = 180/N$. In this work, we have used $N = 127$, shown in figure 3, which gives a precision $\Delta\gamma = 1.42$ degrees. This value is a good compromise between computational resources and the required precision.



(a) $N = 1$ (b) $N = 127$

Figure 3: Comparison of wavelet with parameters $(J, j, N) = (8, 2, 1)$ and $(J, j, N) = (8, 2, 127)$. Both axes are in units of degrees. The wavelet shown in figure (b) is the one used in the analysis.

4. Looking for multiplets

We have developed an algorithm that uses spherical wavelets to identify and locate filaments in maps containing arrival directions of UHECRs. The algorithm is described below.

1. From a set of events, calculate the function $h(\theta, \varphi)$ that represents the signal. This involves choosing a tessellation for the sphere, and counting the total number of events contained in each pixel in the sky.
2. Calculate the Fourier expansion of $h(\theta, \varphi)$, resulting in the coefficients a_{lm} (see equation 4).
3. Choose the appropriate wavelet. For the family of wavelets used, this involves choosing three parameters, the maximum scale J, the scale at which the analysis is performed j and the azimuthal band limit, controlled by the parameter N. Ideally, the angular size of the wavelet will match the size of the multiplet. This choice relies on 1.
4. Calculate $C(\alpha, \beta, \gamma)$ from equation (7).
5. Select all (α, β, γ) such that $|C(\alpha, \beta, \gamma)| > C_0$, where C_0, is threshold value.
6. Select the events at each location (α, β, γ) (see section 4.2). This will result in m groups of events.
7. Discard all groups of events for which $n < n_0$, where n is the number of events in the group, and n_0 is a threshold value.
8. For each group of events, calculate the correlation c of the graph $\delta \times 1/E$ (see equation 17).
9. Accept as a multiplet candidate all groups for which $|c| > c_0$, where c_0 is a threshold value. (see section 4.2)

4.1. Establishing thresholds

To carry out all steps of this algorithm one needs to establish three threshold values: the wavelet threshold C_0 used in step 5, the minimum number of events n_0, used in step 7, and the minimum correlation c_0, used in step 8.

To calculate the thresholds, we have used simulated sky maps containing arrival directions of events isotropically distributed. No magnetic fields are considered in this case, so that if a multiplet is identified by our method, it certainly happened by chance. This allows us to establish the threshold values C_0, c_0, and n_0 by using algorithm previously presented in section 4.

To estimate C_0, we start with M isotropic realizations following the same assumptions on the injection spectrum and exposure. C_i is the largest wavelet coefficient obtained for each realization i, according to step 5 of the algorithm.

By selecting only one value of C for each realization, we have a single correlation coefficient c_i for each of the simulated isotropic skies, calculated at step 8 of the algorithm.

From M wavelet coefficients C_i and correlations c_i, we choose

$$C_0 = \bar{C} + r_C\sigma_C \quad \text{and} \quad c_0 = \bar{c} + r_c\sigma_c, \tag{9}$$

where $\bar{C}$ and $\bar{c}$ denote the average values over M realizations, and σ_C and σ_c their respective standard deviations. The numbers r_C and r_c are chosen according to the error the user is willing to accept. In this work we set $r_C = 1$ and $r_c = -1$.

The threshold value for the number of events in the multiplet, n_0, is chosen depending on the specific problem. In this analysis, for example, we have chosen $n_0 = 10$, following Ref. [4].

4.2. Selecting events

To select events around the position where a wavelet has a high magnitude we need the Euler angles α and β, to provide a location in the sky, and the angle γ, to provide an orientation.

A multiplet can, in principle, describe an arbitrary curve on the surface of the celestial sphere, depending on the intervening magnetic fields. To be general for all possible shapes we have used small segments as shown on figure 4. Each segment should be small enough so that the curve described by the events is approximately a straight line inside the segment.

A small deflection angle ($\lesssim 10°$) can be written as $\delta = \Delta s/r$, which in spherical coordinates takes the form

$$\Delta s^2 = r^2((\Delta\beta)^2 + \sin^2\beta(\Delta\alpha)^2), \tag{10}$$

with Δs being the linear distance over the surface of a sphere of radius r.

Now we note that the unit vectors $\hat{\alpha}$ and $\hat{\beta}$, which point in the direction where α and β vary, are orthogonal and span the plane tangent to the sphere at (α, β). Using equation 10 we can write a displacement vector, tangent to that point as

$$\vec{d} = r(\Delta\beta)\hat{\beta} + r\sin\beta(\Delta\alpha)\hat{\alpha}. \tag{11}$$

It is convenient to write $\vec{d}$ in terms of $\hat{L}$, aligned with the wavelet, and $\hat{W}$, which is perpendicular to it. These vectors are related to $\hat{\alpha}$ and $\hat{\beta}$ by a rotation of an angle γ in the tangent plane, as shown below:

$$\hat{\alpha} = \cos\gamma\hat{L} - \sin\gamma\hat{W} \tag{12}$$
$$\hat{\beta} = \sin\gamma\hat{L} - \cos\gamma\hat{W}. \tag{13}$$

We can now cast 11 as

$$\vec{d} = d_{\hat{L}}\hat{L} + d_{\hat{W}}\hat{W}, \tag{14}$$

where

$$d_{\hat{L}} = r(\Delta\beta\cos\gamma + \sin\beta\Delta\alpha\sin\gamma) \tag{15}$$
$$d_{\hat{W}} = r(-\Delta\beta\sin\gamma + \sin\beta\Delta\alpha\cos\gamma). \tag{16}$$

With equations 15 and 16, we can select events in the sky at the position of maximum wavelet coefficient. The only information required are the Euler angles α, β, and γ, since $r = 1$ for the celestial sphere.

5. Analysis

To simulate both the background and the events belonging to the multiplet, we have used the CRPropa software [32]. Protons were injected following a typical $E^{-2.2}$ spectrum, for $15 < E < 40$ EeV. The effective detector size is a sphere of radius 1 Mpc, and the source was assumed to be in our local universe, at a distance of 30 Mpc from the detector. For the isotropic data sets, the events were simulated in random positions of the sky.

We have divided our analysis in two parts. First we analyze an isotropic distribution of events to establish the thresholds, as explained in section 4.1. In the second part of the analysis we used a simulated multiplet composed by 10 events, embedded in the same isotropic datasets previously used. The multiplet formation was induced by an unrealistic uniform magnetic field of 1 nG oriented in the $\hat{z}$ direction. Nevertheless, the results are totally independent of the method used to generate the multiplets, since it depends only on the position of the events, and on the energy- deflection correlation. The number of isotropic events range between 100 and 1000 events, in steps of 100, and a total of 1000 realizations for each of these values. The multiplet embedded in one of the isotropic skies containing 1000 isotropically distributed events can be seen in figure 5.

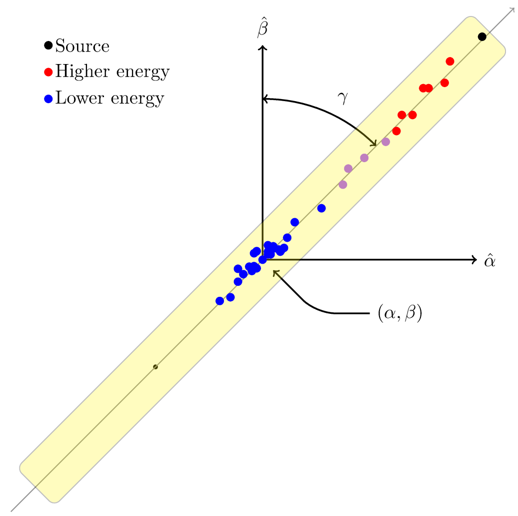


Figure 4: Illustration of a multiplet along a segment. The dots represent events from higher (red) to lower (blue) energies. The angles (α, β, γ) are the Euler angles associated with the multiplet.

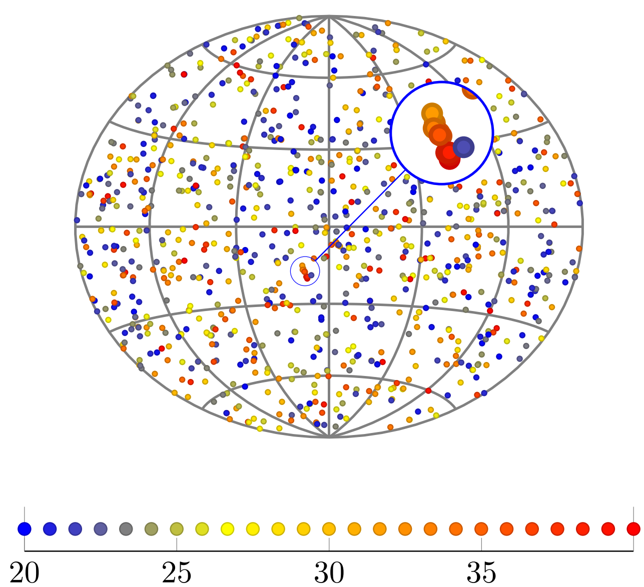


Figure 5: Simulated multiplet containing 10 events embedded in a background of 1000 events. The color scale corresponds to the energy of the events in units of EeV. The zoom shows more closely the events of the simulated multiplet.

The simulated multiplet has a correlation coefficient $c = 0.99$ when no background is present. c corresponds to the Pearson's coefficient of the $\delta \times E^{-1}$ graph, given by:

$$c = \frac{\sum_{i=1}^{n}\left(\frac{1}{E_i} - \left\langle\frac{1}{E}\right\rangle\right)\sum_{i=1}^{n}\left(\delta_i - \langle\delta\rangle\right)}{\sqrt{\sum_{i=1}^{n}\left(\frac{1}{E_i} - \left\langle\frac{1}{E}\right\rangle\right)^2}\sqrt{\sum_{i=1}^{n}\left(\delta_i - \langle\delta\rangle\right)^2}}. \tag{17}$$

The scale at which we performed the analysis was chosen based on table 1. At scales $j = 2$ and 3 the angular size of the wavelet is close to what we would expect for a multiplet similar to the simulated one, i.e., $10° \times 2°$. Since at scale $j = 3$ the maximum precision of the angular variable is $2.8°$, we have used $j = 2$ for the analysis. The size of the segments are $10° \times 2°$, to match the dimensions of the wavelet.

6. Results and Discussion

In this section we present the results of the proposed algorithm, when applied to the simulations.

Since the magnitude of the wavelet coefficients is used to select directions of interest in sky maps, i.e., Euler angles, establishing the wavelet coefficient threshold is a crucial part of the analysis. The wavelets should have enough sensitivity to distinguish the cases of interest (isotropic data sets with an embedded multiplet and without it). Figure 6 shows the average values of the magnitude of the largest wavelet coefficients found in each realization, for the two data sets. The error bars are the standard deviation of the 1000 realizations of isotropic skies. In this particular case, for a multiplet composed of 10 events, the error bars start overlapping at around 1000 events, which corresponds to a fraction of events from the multiplet with respect to the background of $\approx 10^{-2}$.

The correlation coefficient is the main observable to properly characterize a multiplet. In this analysis we have used $n_0 = 10$, which means that a candidate will only be considered a multiplet if it is composed by, at least, 10 events. We also have to determine the value of c_0, which is the value of the threshold correlation coefficient. It is expected that the greater the number of isotropically distributed events, which corresponds to the background, the smaller the correlation coefficient. This is corroborated by figure 7, which shows that the correlation coefficient monotonically decreases for larger number of background events, indicating a contamination of the multiplet by the background. This allows us to establish a threshold value for the correlation coefficient. For 1000 isotropic events the lowest value of c_0 considered is $c_0 = 0.4$, according to equation 9.

It is important to address the questions related to the probability of the algorithm finding a multiplet with $c > c_0$ in a given data set, when no multiplets are known to be present (Type II error), and the probability of a multiplet to have $c < c_0$ if it is known to be present in the data (Type I error). In the isotropic simulations, no multiplets with $c > 0.4$ were identified for a fraction of events inferior to 10^{-2}, implying that the maximum type II error of the method is 10^{-3}. Figure 8 shows the Type I error introduced by our method. It is worth mentioning that for the most conservative case, which has the lowest fraction of number of events of the multiplet with respect to the number of

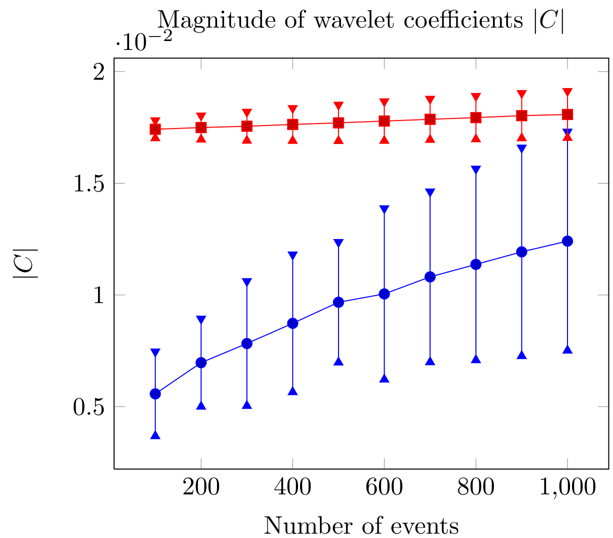


Figure 6: Magnitude of the largest wavelet coefficients in all the 1000 realizations, as a function of the number of events in the background. Blue circles correspond to the purely isotropic case, and red squares to the multiplet embedded in the isotropic background. The error bars are the standard deviation of the wavelet coefficients for all the realizations.

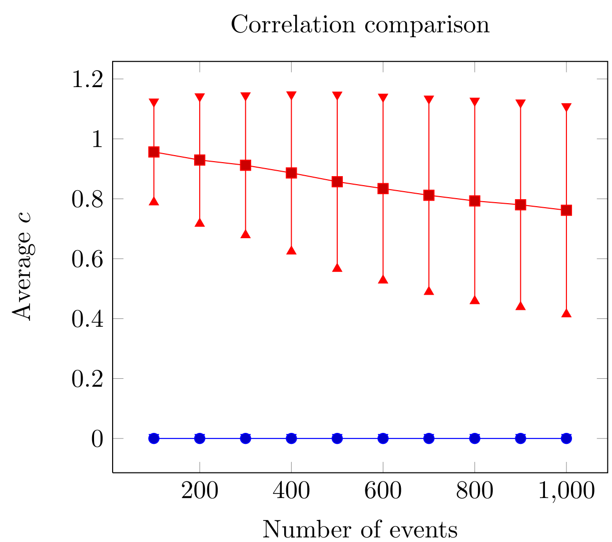


Figure 7: Correlation coefficients of the 1000 realizations, as a function of the number of events in the background. Blue circles correspond to the purely isotropic case, and red squares to the multiplet embedded in the isotropic background. The error bars are the standard deviation of the wavelet coefficients for all the realizations.

events of the background, the Type I error introduced by using a threshold value $c_0 = 0.4$ is approximately 25%, whereas for the least conservative case it is around 2%. For a high threshold value such as $c_0 = 0.9$, this error is of the order of 30% for the most conservative case. It is important to stress that the estimated Type I and II errors depend on the confidence interval adopted, set by the parameters r_C and r_c.

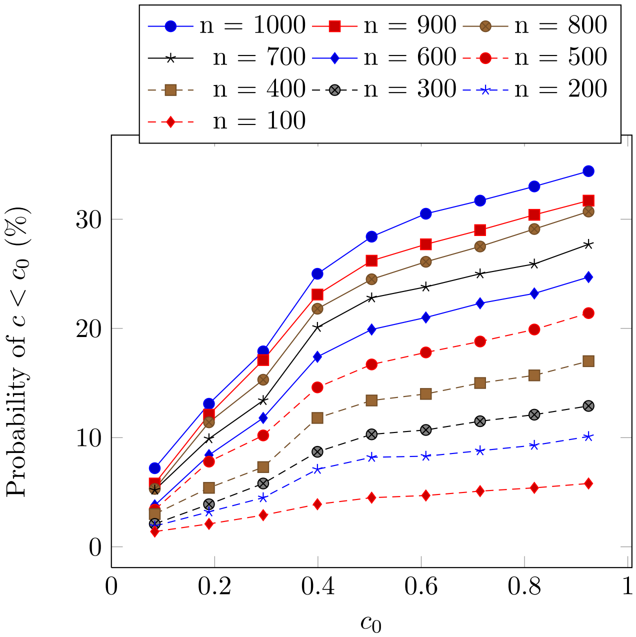


Figure 8: Probability of having $c < c_0$ (type I error) when the multiplet is known to be present in the data set. Each curve corresponds to a different fraction of events of the multiplet with respect to the isotropic background.

7. Conclusion

We have developed a wavelet based analysis method to identify multiplets in maps containing arrival directions of UHECRs. We have illustrated the method by applying it to a hypothetical scenario with a uniform magnetic field of 1 nG, considering several fractions of events from the multiplet with respect to the background, down to a fraction of 10^{-2}. All the parameters used in this study are rather arbitrary. These choices were made to illustrate the method, and not to evaluate its overall performance.

The observables used to accept or reject the multiplet candidate were the magnitude C of the wavelet coefficient at the multiplet location (α, β, γ), the correlation c of the deflection versus inverse of the energy graph and the total number of events in the multiplet n_0.

The efficiency of the method was assessed by comparing the analysis when applied to isotropically distributed events with and without an embedded multiplet. The probability of wrongly accepting a candidate multiplet with a correlation coefficient c above the threshold value $c_0 > 0.4$ is below 10^{-3}. The probability of wrongly rejecting a candidate multiplet, when it is known to be present in the data, is in the most conservative case, approximately 30%. This value goes down to 3% by decreasing the number of isotropic events in the data set. For the adopted correlation coefficient threshold, these values are, respectively, 25% and 2%.

Even though we have assumed a uniform detector exposure for the analysis previously presented, the results also hold for a non uniform exposure, since the ideal parameters for multiplet search in the wavelet analysis suppress low frequency modes, that could be related to the coverage of the cosmic ray detector. The only caveat is that, in this case, the comparison data sets should follow an isotropic distribution modulated by the exposure of the detector.

The method focused on the search of multiplets, but it can be adapted for generic application to related problems in astrophysics, particularly the ones involving the search of filamentary structures in sky maps.

All results showed in this paper can be easily reproduced with the software SWAT (Spherical Wavelet Analysis Tool), developed by the authors of this paper. In case of interest in the code, please contact us.

Acknowledgments

The authors wish to thank the Pierre Auger group of the University of Wuppertal (UW), where part of this work has been written, and part of the simulations done, and specially to Nils Nierstenhöfer for the valuable help with the simulations. We would also like to thank Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) for the financial support through grants 2010/07359-6 and 2010/04743-0. Part of the simulations were performed in the computational system available at Instituto de Física de São Carlos, Universidade de São Paulo (IFSC-USP), granted by FAPESP (2008/04259-0). This project was also supported by Coordenação de Aperfeiçoamento de Nível Superior (CAPES).

References

[1] HiRes Collaboration, Phys. Rev. Lett., 104 (2010) 161101. [arXiv:0910.4184]
[2] Pierre Auger Collaboration, Phys. Rev. Lett. (2010) 104 091101. [arXiv:1002.0699]
[3] D. Harari *et al.*, J. High Energy Phys. 0203 (2002) 45. [astro-ph/0202362]
[4] Pierre Auger Collaboration, Astropart. Phys. 35 (2012) 354. [arXiv:1111.2472]
[5] D. Harari, S. Mollerach, E. Roulet, Astropart. Phys. 25 (2006) 412. [astro-ph/0602153]
[6] G. Golup *et al.*, Astropart. Phys. 32 (2009) 269. [arXiv:0902.1742]
[7] D. Harari, S. Mollerach, E. Roulet, J. High Energy Phys. 8 (1999) 22. [astro-ph/9906309]
[8] X. H. Sun *et al.*, Astron. Astrophys. 477 (2008) 573. [arXiv:0711.1572]
[9] M. S. Pshirkov *et al.*, Astrophys. J. 738 2 (2011) 192. [arXiv:1103.0814]
[10] R. Jansson, G. Farrar, Astrophys. J., 757 (2012) 14. [arXiv:1204.3662]
[11] G. Sigl, F. Miniati, T. Ensslin, Phys. Rev. D 68 (2003) 043002. [astro-ph/0302388]
[12] K. Dolag *et al.*, JCAP 1 (2005) 9. [astro-ph/0410419]
[13] E. Armengaud, G. Sigl, F. Miniati, Phys. Rev. D 72 (2005) 043009. [astro-ph/0412525]

[14] S. Das *et al.*, Astrophys. J. 682 7 (2008) 29. [arXiv:0801.0371]
[15] L. Cayón *et al.*, in de Oliveira-Costa A., Tegmark M., eds, Microwave Foregrounds. ASP Conference Series 181 (1999) 349.
[16] J. L. Sanz *et al.*, Mon. Not. R. Astr. Soc. 309 (1999) 672. [astro-ph/9906367]
[17] J. González-Nuevo *et al.*, Mon. Not. R. Astr. Soc. 369 (2006) 1603. [astro-ph/0604376]
[18] G. Faÿ *et al.*, (2011). [arXiv:1107.5658]
[19] R. Alves Batista *et al.* Physicæ Proceedings 1 (2010) 1-6. [arXiv:1201.2183]
[20] R. Alves Batista, E. Kemp, B. Daniel, Int. J. Mod. Phys. E 20 (2011) 61-66. [arXiv:1201.2799]
[21] A. A. Ivanov, A. D. Krasilnikov, M. I. Pravdin, JETP Lett. 78 (2003) 695.
[22] A. A. Ivanov, A. D. Krasilnikov, M. I. Pravdin, International Cosmic Ray Conference 1 (2003) 341.
[23] Martínez-González E. *et al.*, Mon. Not. R. Astr. Soc. 336 (2002) 22. [astro-ph/0111284]
[24] J. D. McEwen *et al.*, Mon. Not. R. Astr. Soc. 369 (2006) 1858. [astro-ph/0510349]
[25] J. D. McEwen *et al.*, J. Fourier Anal. Appl. 13 (2006) 495. [arXiv:0704.3158]
[26] J. D. McEwen *et al.*, Mon. Not. R. Astr. Soc. 388 (2008) 659. [arXiv:0803.2157]
[27] J. D. McEwen *et al.*, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series 6701 (2007) 670115. [arXiv:0708:3874]
[28] J. D. McEwen *et al.*, Mon. Not. R. Astr. Soc. 384 (2008) 1289. [arXiv:0704.0626]
[29] J. A. Kovacs, W. Wriggers, Acta Crystallogr. D. 58 (2002) 1282.
[30] P. Kostelec, D. Rockmore, J. Fourier Anal. Appl. 14 (2008) 145.
[31] Y. Wiaux *et al.*, Mon. Not. R. Astr. Soc. 388 (2008) 770. [arXiv:0712.3519]
[32] K.-H. Kampert *et al.*, Astropart. Phys. 42 (2013) 41. [arXiv:1206.3132]

Identifying Patterns on Cosmic Ray Maps with Wavelets on the Sphere

Rafael Alves Batista,* Marcelo Zimbres, and Ernesto Kemp
Instituto de Física "Gleb Wataghin" - Universidade Estadual de Campinas

The deflection of ultra-high energy cosmic rays depends on the shape of the injection spectrum of the source and the pervasive cosmic magnetic fields. In this work it is applied the wavelet transform on the sphere to search for energy ordered filamentary structures arisen from magnetic bending. These structures, the so-called multiplets, can bring relevant information concerning the intervening magnetic fields.

I. INTRODUCTION

Ultra-High Energy Cosmic Rays (UHECRs) are particles with energy above 10^{18} eV reaching the Earth. Almost a century after the discovery of cosmic radiation, the origin, chemical composition and mechanisms of acceleration and propagation of the UHECRs remain a mystery[1].

The way their arrival directions are represented is by associating each event with its corresponding latitude and longitude in the celestial sphere. For charged particles, the pervasive cosmic magnetic fields play an important role, affecting the distribution of arrival directions of UHECRs.

Several mechanisms of acceleration of UHECR in different astrophysical sites result on a power law differential energy spectrum[3], i. e.,

$$\frac{dN}{dE} \propto E^{-\alpha}, \tag{1}$$

where α is the spectral index. For a single source emitting UHECRs of energy E according to this spectrum, the expected angular deflection is inversely proportional to the energy. Therefore, particles with different energies will deflect differently and the greater the energy of the particle the smaller the angular deflection of the particle (and vice-versa). So, if locally, the typical scale of the intervening magnetic fields (coherence length) is smaller or of the same order of magnitude of the Larmor radius of the cosmic ray, filamentary structures in sky maps are expected to be formed. The cosmic ray events that compose this structure are ordered by energy, creating the so-called multiplets of UHECRs. For the galactic magnetic field the expected angular deflection (δ) is given by[2]

$$\delta = 8.1^{\circ} 40 \frac{Z}{E} \left| \int_0^L \frac{d\vec{r}}{3\ \text{kpc}} \times \frac{\vec{B}}{2\ \text{kpc}} \right|, \tag{2}$$

where L is the distance to the source, $\vec{B}$ is the magnetic field, Z is the atomic number of the nuclei and E its energy in EeV.

II. THE METHOD

In this work it is proposed a new method (for another method refer to [4]) to identify multiplets in maps containing arrival directions of UHECRs. This method is based on the spherical wavelet transform[5] and consists on the correlation of a signal with rotated versions of a given pattern, both defined on a spherical manifold. Let a_{lm} be the coefficients of the spherical harmonic expansion of the signal, and b_{ln} be the same expansion for the sought rotated pattern (in this case, a wavelet). The correlation coefficient is

$$C = \sum_{l=0}^{B-1} \sum_{|m| \le l} \sum_{|m'| \le l} \overline{a_{lm}} b_{lm'} D^l_{mm'}(\alpha, \beta, \gamma), \tag{3}$$

where the overline indicates the complex conjugate, $D^l_{mm'}$ designates Wigner-D functions, B is the band limit of the implemented algorithm, and (α, β, γ) are the Euler angles associated to the rotation of the wavelet.

Applying the spherical wavelet transform using a code called SWAT (Spherical Wavelet Analysis Tool)[6] to the distribution of arrival directions of UHECRs, if the correlation coefficient is high, this indicates a similarity between the pattern (wavelet) used and the signal. The Euler angles provide information regarding the coordinates (latitude and longitude) where there is a maximal correlation coefficient, and the orientation of the pattern.

Even though the method is able to identify filamentary structures, it does not take into account the energy ordering of the events for the analysis. This is done by choosing a rectangle of dimensions $\sim 12^{\circ} \times 1.5^{\circ}$ (typical dimensions of a multiplet) with the obtained orientation, and calculating the Pearson coefficient for $\delta \times E^{-1}$ of the events within this rectangle. If this coefficient is high, it indicates an energy ordering and the filamentary structure is a possible multiplet candidate.

III. APPLICATION TO SIMULATIONS

Events of UHECRs were simulated using the code CRT[7], which numerically solves Lorentz force equations of motion for charged particles in a magnetic field using fifth order adaptative Runge-Kutta routines. Fifty proton events were simulated from a hypothetical source lying at a distance of 4 kpc from Earth, with galactic coordinates $(l, b) = (90^{\circ}, -45^{\circ})$. The source was assumed to have an injection spectrum with $\alpha = 2.7$ and to emit particles within a perfectly collimated jet.

The magnetic field models used in the simulation are the ones proposed by Harari, Mollerach and Roulet[8], namely the ASS-S, ASS-A, BSS-S and BSS-A models, were ASS stands for AxiSymmetric Spiral (even under the transformation $\theta \to \theta + \pi$, where θ is the azimuth angle) and BSS for BiSymmetric Spiral (odd with respect to the mentioned transformation). The sufixes -A and -S indicate the parity of the

* rab@ifi.unicamp.br

transformation above and below the galactic disk, where -S is symmetric (even) and -A is antisymmetric (odd) with respect to the $z \to z$ transformation. In this class of models the magnetic field in the disk of the galaxy can be decomposed in a radial and an azimuthal part, depending on the pitch angle (p) of the logarithmic spiral that models the magnetic field of the Milky Way. The intensity of the magnetic field in the disk is

$$B(r,\theta) = 3\frac{r_0}{r}\tanh^3\left(\frac{r}{r_1}\right)\cos^m\left(\theta - \frac{1}{\tan p}\ln\frac{r}{r_0}\right)\ \mu G, \quad (4)$$

where $r_0 = 10.55$ kpc, $r_1 = 2$ kpc, $p = -10^\circ$ and m is 1 for BSS models and 2 for ASS models. The z component of the magnetic field is

$$B(z) = \left[2\cosh\left(\frac{z}{z_0}\right) + 2\cosh\left(\frac{z}{z_1}\right)\right]^{-1}\zeta(z)\ \ \mu G, \quad (5)$$

where $\zeta(z) = \tanh(z/z_2)$ for -A models and 1 for -S models, $z_0 = 4$ kpc, $z_1 = 0.3$ kpc and $z_2 = 20$ pc.

By applying the method to these specific simulations, using the wavelet transform with suitable parameters (compatible with the expected dimensions of a multiplet), the multiplet was correctly identified and the orientation for different models was calculated and is shown in table I. Figure 1 shows the simulated events for the simulations.

Tabela I: Orientation (χ) for the simulated source in the class of models of Harari, Mollerach and Roulet, for the different symmetries.

Model	χ (°)
ASS-S	100.3
ASS-A	108.4
BSS-S	65.7
BSS-A	68.1

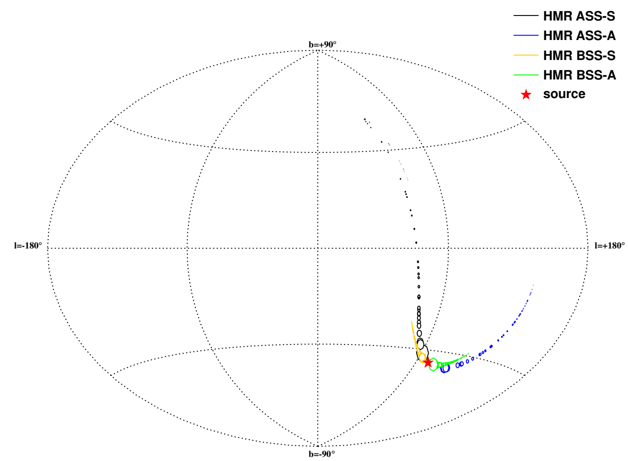


Figura 1: Skymap containing the arrival direction of the simulated events. The size of the circles are proportional to the energy of the event.

IV. CONCLUSIONS AND PERSPECTIVES

The orientation of the multiplet, i.e., the angle it forms with the galactic plane, depends on the coordinates of the source and the magnetic field model adopted. Therefore, it is possible to set limits on models of the galactic magnetic field by analysing the orientation of multiplets. When the method was applied to simulated data, the multiplets were correctly identified and it was noticed a relation between the orientation of the multiplet and the adopted model. If this method is applied to data collected by the Pierre Auger Observatory or any other cosmic ray experiment and is able to successfully identify multiplets, a new tool to probe the galactic magnetic field will be available.

ACKNOWLEDGEMENTS

We are grateful for the financial support of FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) and CAPES (Coordedoria de Aperfeiçoamento de Pessoal de Nível Superior).

[1] M. Nagano and A. A. Watson, Rev. Mod. Phys., 72:689, 2000.
[2] G. Giacinti, X. Derkx, and D. V. Semikoz, JHEP, 3:22, 2010.
[3] R. J. Protheroe. In: Topics in Cosmic-Ray Astrophysics, 247, 1999.
[4] Pierre Auger Collaboration. Astropart. Phys. 35(6):354, 2012.
[5] Y. Wiaux *et al.*, MNRAS, 388:770, 2008.
[6] Available in `www.ifi.unicamp.br/~mzimbres`
[7] M. S. Sutherland, B. M. Baughman, and J. J Beatty. Astr. Phys., 34:198, 2010.
[8] D. Harari, S. Mollerach, and E. Roulet, JHEP, 8:22, 1999.

Referências Bibliográficas

[1] V. Hess, Physik. Zeitschr. **13** (1912) 1084 - 1091.

[2] N. Chiba, *et. al.*, Nucl. Instrum. Methods A **311** (1992) 338 - 349.

[3] H. Ohoka, *et. al.*, Nucl. Instrum. Methods A **385** (1997) 268 - 276.

[4] G. T. Zatsepin, V. A. Kuz'min, Journal of Experimental and Theoretical Physics Letters 4: 78–80.

[5] T. Stanev, *High Energy Cosmic rays.* (Praxis Publishing Ltd. Chichester, UK, 2004).

[6] R. U. Abbasi, *et. al.*, Phys. Rev. Letters. **92** (2004) 151101.

[7] J. Abraham, *et. al.*, Nucl. Instrum. Methods A **523** (2004) 50 - 95.

[8] N. W. Boggess, *et. al.*, 21st General Assembly Of The International Astronomical Union. Buenos Aires, Argentina, Jul 23-01, 1991

[9] C. L. Bennett, M. Halpern, G. Hinshaw, *et. al.*, Astrophysical Journal Supplement, Vol. 148 Issue: 1, Sep 2003.

[10] Pierre Auger Collaboration, Astropart. Phys. 35 (2012) 354.

[11] T. K. Gaisser, *Cosmic rays and particle physics* (Cambridge University Press, Cambridge; New York, 1990).

[12] G. Golup *et al.*, Astropart. Phys. 32 (2009) 269.

[13] Y. Wiaux *et al.*, Mon. Not. R. Astr. Soc. 388 (2008) 770.

[14] HiRes Collaboration, Phys. Rev. Lett., 104 (2010) 161101.

[15] D. Harari *et al.*, J. High Energy Phys. 0203 (2002) 45.

[16] R. U. Abbasi *et al.*, Proceedings of the 30th International Cosmic Ray Conference (2007).

[17] Trapani S, Navaza J., Acta Crystallogr A. **62** (2006) 262 - 9.

[18] `http://www.fftw.org`

[19] Advances in Applied Mathematics **15**, 202-250 (1994)