

CENTRO NACIONAL DE PESQUISA EM ENERGIA E MATERIAIS |  
CNPEM  
ILUM - ESCOLA DE CIÊNCIA

RELATÓRIO FINAL  
INICIAÇÃO À PESQUISA III

**MÉTODOS DE APRENDIZADO DE MÁQUINA A PARTIR  
DE ESPECTROS DE RAIOS X USANDO SIMULAÇÕES E  
DESCRIPTORES**

**MACHINE LEARNING METHODS BASED ON X-RAY  
ABSORPTION SPECTRA USING SIMULATIONS AND  
DESCRIPTORS**

---

**CAUÊ GOMES CORREIA DOS SANTOS**

CENTRO NACIONAL DE PESQUISA EM ENERGIA E MATERIAIS - CNPEM | ILUM - ESCOLA DE CIÊNCIA  
RUA LAURO VANUCCI, 1020 - FAZENDA SANTA CÂNDIDA, CAMPINAS - SP, 13087-548

**EMELYN ALVES**

CENTRO NACIONAL DE PESQUISA EM ENERGIA E MATERIAIS - CNPEM | ILUM - ESCOLA DE CIÊNCIA  
RUA LAURO VANUCCI, 1020 - FAZENDA SANTA CÂNDIDA, CAMPINAS - SP, 13087-548

**JAMES MORAES DE ALMEIDA**

CENTRO NACIONAL DE PESQUISA EM ENERGIA E MATERIAIS - CNPEM | ILUM - ESCOLA DE CIÊNCIA  
RUA LAURO VANUCCI, 1020 - FAZENDA SANTA CÂNDIDA, CAMPINAS - SP, 13087-548

**SANTIGO JOSE ALEJANDRO FIGUEROA**

CENTRO NACIONAL DE PESQUISA EM ENERGIA E MATERIAIS - CNPEM | LABORATÓRIO NACIONAL  
DE LUZ SÍNCROTRON  
POLO II DE ALTA TECNOLOGIA - R. GIUSEPPE MÁXIMO SCOLFARO, 10000 - BOSQUE DAS PAL-  
MEIRAS, CAMPINAS - SP, 13083-100

30 de novembro de 2024  
Campinas - SP | Brasil

## Resumo

### Implementação de método para determinação de máximos e mínimos em espectros XANES

Dentre as técnicas que serão disponibilizadas pela linha de luz QUATI (Quick X-ray Absorption Spectroscopy for Time and space resolved experiments), destaca-se o XANES (X-ray Absorption Near Edge Structure), uma técnica fundamental para estudos da estrutura eletrônica de materiais. No entanto, os espectros XANES são de difícil interpretação devido a sua natureza não analítica, o que demanda métodos automáticos para a identificação de picos característicos. Dessa forma, este trabalho visa implementar o método Rising-Sun, descrito no artigo *“The Rising Sun Envelope Method: an automatic and accurate peak location technique for XANES measurements”*, para a identificação automática e precisa de máximos e mínimos nos espectros XANES. A implementação do método Rising-Sun foi realizada para análise de arquivos ‘.xdi’ da Cruzeiro do Sul Utils, biblioteca da base de dados Cruzeiro do Sul, que reúne diversos espectros de XAS (X-ray Absorption Structure). Para tal, utilizou-se linguagem Python de programação e arquivos jupyter notebook. O objetivo é facilitar a comparação de espectros e auxiliar futuros usuários da linha QUATI na interpretação rápida e precisa de seus dados, aumentando a eficiência em estudos de materiais. A implementação do método Rising-Sun permitiu a identificação de picos e vales em espectros XANES disponíveis na Cruzeiro do Sul Utils. Com tais informações, a possibilidade de comparação de espectros XAS a partir de métodos de aprendizado de máquina, que utilizem simulações e descritores, torna-se mais factível em futuras aplicações.

## Implementação de método para determinação estados de oxidação em espectros XANES

Por carregar muitas informações sobre a sua amostra, o espectro de XANES pode ser utilizado em diversas técnicas, objetivando o entendimento de diferentes características do material. Uma grandeza importante para caracterizar materiais e compreender melhor o ambiente químico que os compostos estão envoltos é o estado de oxidação (ou número de oxidação). E como o XANES está intimamente ligado ao estudo da estrutura eletrônica, é possível inferir e estudar o estado de oxidação de elementos partindo de seu espectro de XANES. Logo, este trabalho também visa implementar um método de estimar o estado de oxidação, partindo de arquivos '.xdi' da Cruzeiro do Sul Database. Esse método se baseia em trabalhos desenvolvidos anteriormente na linha e no trabalho *“X-ray-absorption edge shifts in rare-earth-transition-metal compounds”*[1]. Esta parte do trabalho objetiva obter uma maneira rápida de comparação de estados de oxidação esperados de maneira precisa, permitindo que os futuros usuários da linha QUATI maior eficiência e qualidade de tempo de estudo na linha de luz. De maneira análoga a outra sessão do presente trabalho, a possibilidade de identificação de estados de oxidação de maneira rápida abre portas para a utilização de métodos de aprendizado de máquina, de maneira a auxiliar ainda mais os pesquisadores e usuários futuros.

**Palavras-chave:** XANES, Identificação de picos e vales, Método Rising-Sun, Estado de Oxidação, Interpretação de espectros, Aprendizado de máquina, Linha de Luz QUATI, Cruzeiro do Sul Utils.

## Resumo das atividades realizadas no período

Primeiramente, os fundamentos básicos teóricos a respeito de XAFS e XANES foram adquiridos a partir dos dois primeiros capítulos do livro “Fundamentals of XAFS” de Matthew Newville.

Todos os códigos desenvolvidos no projeto estão em linguagem de programação Python e foram executados no formato Jupyter Notebook. Os arquivos de código dos métodos implementados estão disponíveis em: [https://github.com/CaueSantos1812/RisingSun\\_Oxidation](https://github.com/CaueSantos1812/RisingSun_Oxidation).

Os arquivos '.xdi' foram obtidos da Cruzeiro do Sul Database e processados pela biblioteca Cruzeiro do Sul Utils. Os métodos para determinação de máximos e mínimos e estados de oxidação foram implementados, separadamente. Os resultados de ambos foram obtidos ao longo do período de implementação.

Após análise de resultados, foi possível desenvolver o relatório e apresentação finais.

# 1 Introdução

## 1.1 Introdução: XANES [2, 3]

*XANES* na verdade é uma técnica que explora como a probabilidade de absorção de raios-x por um átomo varia em função da energia, especialmente próximo e acima das energias de ligação dos elétrons de níveis mais internos, conhecidos como elétrons de caroço. Especificamente, o XANES é feito nas regiões próximas da borda de absorção, permitindo uma análise mais detalhada da densidade de estados disponíveis, estado eletrônico, estados de oxidação e ambiente químico próximo ao átomo. As possibilidades de análise e as forças da técnica estão intimamente ligada com a natureza física por trás do XANES, que é quem permite que tantas grandezas e características sejam extraídas do espectro de absorção.

Isso porque, por trás do XANES está o estudo de transições eletrônicas específicas, encadeadas pelos fótons de raios X, responsáveis por excitar elétrons de orbitais internos ( $1s$  ou  $2p$ ) para estados eletrônicos desocupados acima do nível de Fermi daquele átomo[3]. Entretanto, nem todo nível eletrônico participará da absorção de raios X, isto porque para isso acontecer a energia de ligação deste estado deve ser menor que a energia do raio X incidente. Caso contrário, o elétron não será perturbado. E, se caso a energia incidente seja maior do que a energia do estado eletrônico, o raio X será absorvido e promoverá um elétron a um estado de maior energia, e qualquer energia excedente será dada a um fóton-elétron que será emitido do átomo. O fenômeno por trás está intimamente ligado ao *Efeito fotoelétrico*, outro fenômeno físico, que inclusive rendeu ao Einstein o seu prêmio Nobel.

Essas transições são governadas por diversas regras de seleção, além da energia envolvida, o que implica que os estados finais observados no espectro estão relacionados com a densidade local de estados e a natureza quântica das interações atômicas. Com isso características do ambiente químico podem afetar consideravelmente o espectro obtido, já que por exemplo, hibridizações entre os orbitais de um elemento com átomos vizinhos podem dar origem a novas

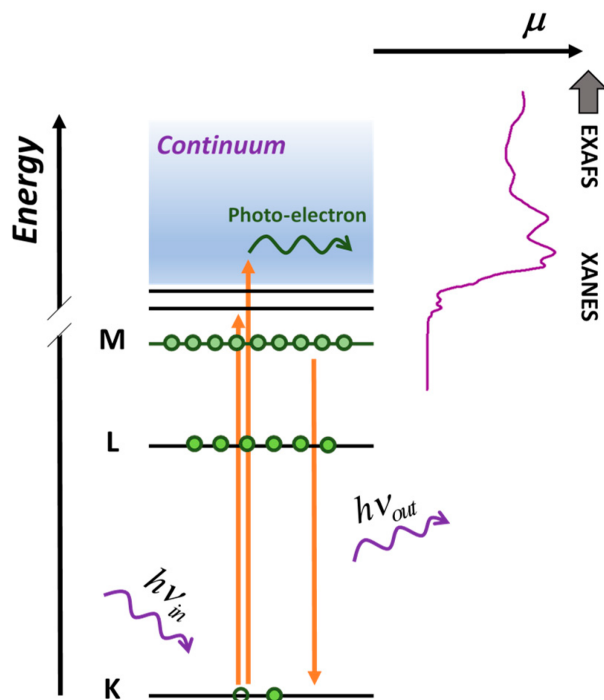


Figura 1: Imagem exemplificando o processo de excitação do elétron[2]

transições, permitindo assim modificações no espectro, como por exemplo a mudança na posição da borda de absorção e a existência de pré-picos.

Assim, a técnica XANES são dependentes da variação da energia do fóton de raios X incidido na amostra, ou seja, são funções da energia ( $\epsilon$ ). Porém o que obtemos não é a energia do foton-elétron, por exemplo, mas sim um coeficiente de absorção ( $\mu$ ) que dá a probabilidade que o raio X vai ser absorvido, de acordo com a lei de *Lambert*[3], que diz que:

$$I = I_0 e^{-\mu t}$$

Na qual  $I$  é a intensidade do feixe que passou pela amostra,  $I_0$  é a intensidade do feixe incidido e  $t$  é a espessura da amostra

É claro que para ser o observável de XANES,  $\mu$  possui uma dependência da energia, porém existe uma outra dependência nessa função que também é importante para a técnica. Que é a dependência do número atômico ( $Z$ ), na verdade é possível estimar  $\mu$ , e por esse cálculo se encontra que  $\mu \propto Z^4$  o que permite afirmar que os valores de espectros de elementos próximos serão suficientemente distintos. Por fim o XANES pode ser feito de duas maneiras, as

quais possuem duas formas diferentes de relacionar  $\mu$  com  $I$  e  $I_0$ : Absorção e Fluorescência com as relações sendo:

$$\mu_{abs}(\varepsilon) = \log\left(\frac{I_0}{I}\right)$$

$$\mu_{flu}(\varepsilon) \propto \frac{I_f}{I_0}$$

Dessa forma, o XANES apresenta ser uma técnica muito avançada e adequada para o estudo de materiais e compostos, com a característica de ser sensível a mudanças no elemento analisado e as condições das proximidades químicas desse átomo. E com o intuito de extrair essas informações que a técnica carrega consigo, diversos métodos de análise de espectro foram desenvolvidos com o passar dos anos, permitindo o entendimento de diversas grandezas a partir do XANES.

Neste trabalho, utilizamos o XANES para investigar: 1. o estado de oxidação de diferentes elementos por meio de uma análise detalhada da variação de energia e 2. determinar os máximos e mínimos do espectro a partir de uma análise matemática. Os métodos permitem, em conjunto, relacionar mudanças na borda de absorção e na densidade de estados disponíveis com propriedades químicas e eletrônicas dos átomos analisados, e possibilitar a utilização de máximos e mínimos como *fingerprints*.

## 2 Experimental (Materiais e Métodos)

### 2.1 Implementação de método para determinação de máximos e mínimos em espectros XANES

O projeto foi desenvolvido utilizando a linguagem de programação Python. O *script* contendo todos os passos para implementação do Método Rising Sun foi executado no ambiente Jupyter Notebook. O código do método Rising Sun foi obtido do repositório no GitHub de Rafael Monteiro, um de seus desenvolvedores. Arquivos de dados experimentais no formato '.xdi', contendo informações sobre os espectros analisados, foram extraídos da Cruzeiro do Sul



Database.

Foi realizada uma curadoria inicial para retirar arquivos duplicados ou corrompidos. Os arquivos selecionados para análise foram processados pela biblioteca Cruzeiro do Sul Utils (CZDS Utils), responsável pela leitura e normalização dos espectros. Após esse processamento, os arquivos '.xdi' foram convertidos em arquivos '.csv' contendo dados de energia (eV) e intensidade de absorção. As primeiras e segundas derivadas dos espectros foram calculadas pela função `calculate_derivatives()`.

Em seguida, o método Rising Sun foi aplicado para identificar os máximos e mínimos de cada espectro. Esse processamento foi realizado submetendo o script como um job no Heisenberg, o HPC da Ilum. Para os espectros que apresentaram pelo menos 10 picos identificados, foi gerado um arquivo (.csv) contendo as informações de energia (eV) e intensidade de absorção dos 10 primeiros picos.

Por fim, a análise de componentes principais (PCA) foi aplicada para avaliar o comportamento dos dados em função do elemento principal do dado experimental. [4]

**XAFS Data Interchange Format (XDI):** Formato de arquivo que contém informações fundamentais para compreender e trabalhar com espectro de absorção de raio-x. Os arquivos são específicos para cada medida e espectros.[5]

**Cruzeiro do Sul Database:** Base de dados que reúne mais de 300 arquivos '.xdi' de espectroscopia de absorção de raio-x (XAS) para deixá-los acessíveis aos futuros usuários da linha de luz QUATI. [4]

**Cálculo de derivadas:** A função '`calculate_derivatives()`' calcula a primeira e a segunda derivadas dos arquivos '.csv' "pré-implementação", salvando os resultados em arquivos separados no mesmo formato. É importante destacar que os valores de primeiras e segundas derivadas computados não são utilizados para determinação de máximos e mínimos pelo método Rising Sun. No entanto, são valiosos para outras análises, como a determinação da borda de absorção, que pode ser identificada pelo primeiro ponto de inflexão, correspondendo ao ponto onde a segunda derivada da função é nula.[4][6]



## Método Rising Sun:

O método Rising-Sun é uma abordagem automática, fundamentada no Teorema de Rolle e no Lema do Sol Nascente de Riesz, para a identificação de máximos e mínimos em espectros. Ele combina o uso de parâmetros limiares, que preservam a localização e a intensidade dos picos, com operadores matemáticos, responsáveis por reduzir o ruído nos espectros. Essa integração garante maior controle e precisão no processo de determinação dos picos.

De modo geral, o parâmetro limiar de altura ( $h$ ) é utilizado para distinguir picos reais de flutuações aleatórias de ruído, considerando um pico significativo aquele cuja altura excede esse limiar. Já o parâmetro limiar de distância ( $d$ ) define a separação mínima entre picos consecutivos, auxiliando na identificação de picos muito próximos, que poderiam ser confundidos como um só. Ambos os limiares devem ser ajustados conforme as características de cada espectro ou espectroscopia. <sup>1</sup>

Os operadores *Rising Sun* (Sol Nascente) e *Valley of Shadow* (Vale das Sombras) são funções auxiliares responsáveis por regularizar o espectro de XANES e calcular a diferença entre a função original, respectivamente. <sup>2</sup> O primeiro, ao ser aplicado sucessivamente em diferentes intervalos de energia, garante que o primeiro pico localizado permaneça no mesmo ponto em cada intervalo, preservando, assim, tanto a localização quanto a intensidade dos picos. O operador Rising Sun é dado por:

$$R_{\mu}(E) = \max_{e_{-\infty} \leq x \leq E} \mu(x)$$

Enquanto o Valley of Shadow, é definido por:

$$V_{\mu}(E) = R_{\mu}(E) - \mu(E)$$

A detecção de máximos e mínimos é acompanhada por uma classificação da “qualidade do pico”, determinada pela “relevância do pico”. Esta, por sua vez,

<sup>1</sup>Verificar secção *Definition 1.1 (Peak Criteria with Thresholds ( $h_*$ ,  $d_*$ ))* do artigo [referência Rising Sun] para mais detalhes.

<sup>2</sup>Verificar secção *Definition 1.3 Rising Sun and Valley of Shadows* do artigo [referência Rising Sun] para mais detalhes.

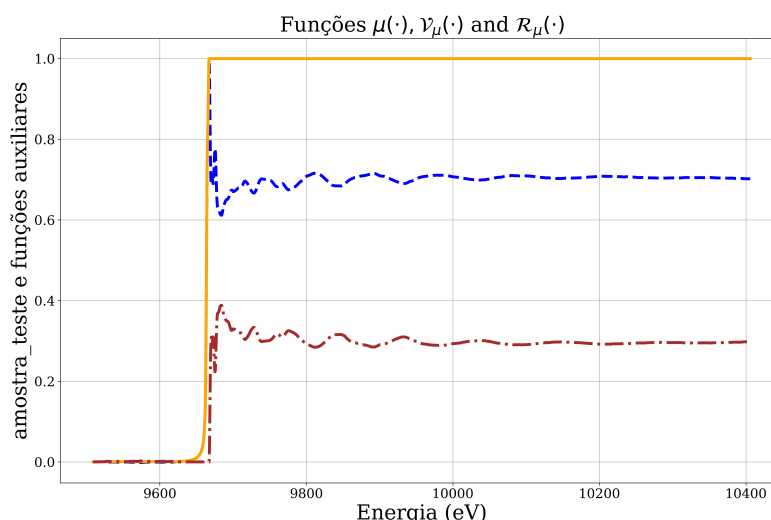


Figura 2: Gráfico de energia por intensidade de absorção com 3 funções plotadas: 1. função original do espectro de absorção de raio-x azul (tracejado); 2. operador Rising Sun em amarelo (linha contínua); 3. operador Valley of Shadows em marrom (tracejado com pontos).

é calculada a partir da razão entre a oscilação do espectro e o *jump* (mudanças abruptas de intensidade entre picos consecutivos) entre o pico atual e o anterior. O critério de classificação é a seguinte: a) Verde para razão menor que 1; b) Amarelo para razão entre 1 e 2; c) Vermelho para razão superior a 3.[6]

**Outros pontos:** O código numérico do Rising Sun passou por uma série de ajustes para permitir a análise de um espectro por vez, uma vez que o método foi originalmente desenvolvido para análise de múltiplos espectros simultaneamente. Além disso, o código foi utilizado até a etapa de identificação dos picos, sendo que a última função aplicada foi a `'plot_marked_peak_spectrum()'`, responsável por plotar o espectro com os picos identificados.

**Heisenberg, o HPC:** A utilização do ambiente do HPC permitiu a execução do código de forma mais eficiente. O tempo necessário para identificar os picos em um espectro, que inicialmente era de aproximadamente 5 minutos, foi reduzido para cerca de 1 minuto no HPC. Essa redução permitiu processar iterativamente todos os arquivos `'xdi'` selecionados, otimizando o tempo de análise.

## 2.2 Implementação de método para determinação estados de oxidação em espectros XANES

Para a implementação do método para determinação do estado de oxidação de compostos “3d”, utilizamos ferramentas computacionais e técnicas de análise. Entre elas: scripts na linguagem Python, notebooks interativos no Jupyter Notebook e a execução de cálculos em jobs submetidos ao HPC da Ilum. Essa abordagem permitiu um processamento eficiente e a manipulação de grandes volumes de dados espectroscópicos, visto que estávamos lidando com a base de dados Cruzeiro do Sul. O método implementado foi desenvolvido com base em dois artigos bem conhecidos e importantes da literatura de XANES, especialmente se tratando da análise do estado de oxidação[7, 1].

Em seu artigo, Wong [7] estuda os espectros XANES de diversos óxidos de vanádio, especificamente olhando para a borda de absorção K, configurando uma análise de transições eletrônicas de caroço e conclui que existe uma relação linear entre o número de oxidação e o *shift* da borda de absorção, isto é, quanto maior for o estado de oxidação, mais a borda estará deslocada para uma região de maior energia [7]. Posteriormente Capehart estudou o mesmo fenômeno de *shifts* na borda de absorção K, mas agora em metais de transição terra-rara [1]. Neste artigo o estudo foi generalizado para uma relação matemática mais rigorosa do que a encontrada por Wong[7]. Esta relação encontrada no artigo de Capehart é o método que foi modificado pelo Santiago e o pessoal da linha [8], e esse sim, foi implementado em python neste trabalho, utilizando os espectros encontrados pela biblioteca Cruzeiro do Sul Utils.

Esse método utiliza uma abordagem quantitativa baseada na integral da borda de absorção no espectro XANES. A ideia central é calcular a área abaixo da curva do coeficiente de absorção ( $\mu(\varepsilon)$ ) em um intervalo de energia delimitado ao redor da borda de absorção K.

$$\sigma_a(\varepsilon_a) = \int_{-\infty}^{\varepsilon_a} (\mu(\varepsilon) + 1) d\varepsilon [1]$$

$$\sigma_r(\varepsilon_r) = \int_{-\infty}^{\varepsilon_r} (\mu(\varepsilon) + 1) d\varepsilon [1]$$

Entretanto, o que vincula o valor de referência e o espectro analisado, é que os dois devem ter o mesmo valor dessa integral, ou seja, o valor debaixo da curva dos dois espectros deve ser a mesma. Por isso, juntando com a relação linear de Wong [7], existirão dois valores diferentes de energia que delimitam o final da integral do espectro de referencia e do espectro analisado. Com isso, se  $\sigma_a = \sigma_0$ , então  $\Delta E$  pode ser calculado como  $\Delta E = \varepsilon_a - \varepsilon_0$  [1].

Esse foi o método implementado nesse trabalho para poder encontrar o estado de oxidação, e como foi dito, sua implementação foi feita a partir de códigos em python que realizavam este cálculo numericamente. Isso foi realizado por meio de um código já existente na linha QUATI [8], que teve que ser adaptado para poder realizar os cálculos com arquivos '.xdi' como input. Uma vez com o código adaptado, foi feita uma etapa de separação de espectros adequados para a análise com este método, ou seja, espectro de compostos 3d. Outra característica desejada é que o estado de oxidação pudesse já ser informado pelo nome do arquivo, ou então que fosse possível de ser calculado por meio da estequiometria da amostra. Isso foi realizado utilizando a biblioteca Cruzeiro do Sul Utils e um script em python, que abria cada um dos XDIs da base de dados, e procurava se o arquivo atendia as condições necessárias.

Isso foi importante pois o próximo passo foi realizar uma reta de calibração para cada elemento analisado, permitindo uma função de  $\Delta E$  que retorna um estado de oxidação estimado. Isso foi feito por meio dos pontos conhecidos dos espectros calculados, por meio de coordenadas de  $(\Delta E, NOX)$  e posteriormente uma regressão linear simples.

### 3 Resultados e Discussão

#### 3.1 Implementação de método para determinação de máximos e mínimos em espectros XANES

A implementação do método Rising-Sun permitiu a identificação da maioria de máximos e mínimos de grande parte dos espectros analisados, como é possível observar a seguir:

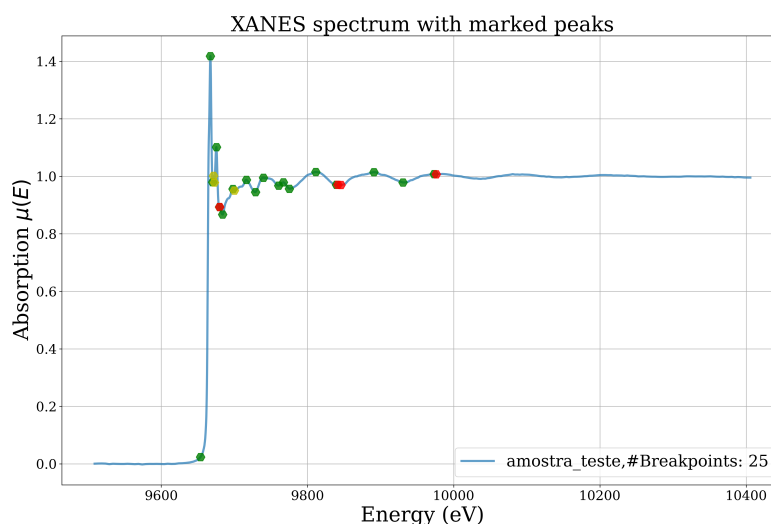


Figura 3: Espectro da Cruzeiro do Sul Database com picos identificados pelo Método Rising Sun

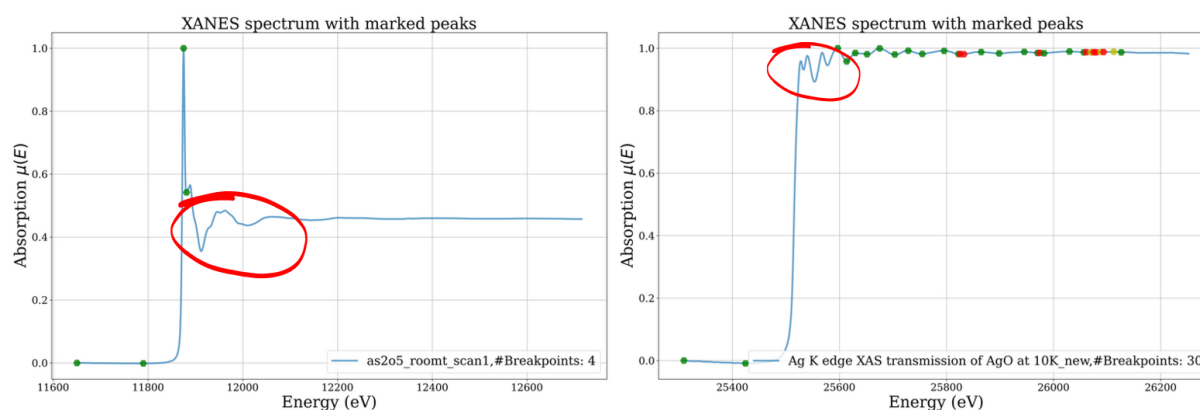


Figura 4: Casos patológicos. As áreas circuladas em vermelho indicam os locais onde não houve identificação de picos e vales.

No entanto, esse padrão de identificação não foi observado em todos os espectros. Como são os casos apresentados na figura 5, nas quais alguns máximos e mínimos não foram determinados antes do primeiro máximo (à esquerda) e alguns deixaram de ser identificados (à direita), mesmo após a identificação de alguns pontos.

A não identificação de pontos de máximos e mínimos é relacionada com a definição dos parâmetros limiares iniciais,  $h^*$  e  $d^*$ , de acordo com Rafael Monteiro. Isso porque esses parâmetros devem ser ajustados levando em consideração o espectro analisado e a técnica de espectroscopia, como já foi comentado. A “generalização” desses parâmetros pode ter prejudicado a identificação de picos e vales pelo método.

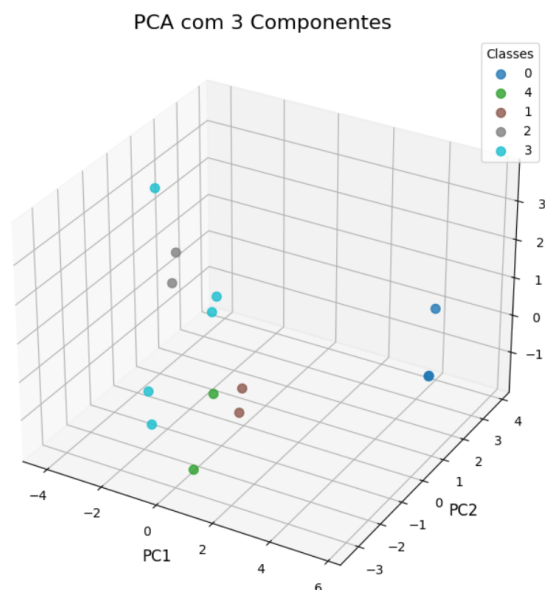


Figura 5: Análise de PCA com 3 componentes.

A partir do PCA é possível verificar que os espectros com elemento principal iguais ficaram mais próximos. Cada classe corresponde a um elemento principal: 0-Prata (Ag), 1-Bromo (Br), 2-Cobalto (Co), 3-Ferro (Fe) e 4-Arsênio (As). De todas as classes a única que apresenta mais esparsidade é a 3, mas também é a que mais tem exemplos.

### 3.2 Implementação de método para determinação estados de oxidação em espectros XANES

Com os estados de oxidação determinados e os  $\Delta E$ 's calculados para todos os arquivos XDI selecionados (e organizados por elementos), foi possível realizar uma regressão linear usando os pontos encontrados. Um exemplo de regressão linear encontrada para o manganês:

Dessa forma, implementando o método da linha QUATI [8], inspirado em Capehart [1], conseguimos encontrar a relação linear que Wong definia [7][9], para todos os compostos "3d" presentes na Cruzeiro do Sul. Entretanto, algumas limitações foram encontradas durante o percurso. Isto porque para elementos que possuem muitas amostras de um mesmo número de oxidação, existe uma dispersão considerável dos valores de  $\Delta E$  referentes ao mesmo NOX, o que diminui a qualidade da regressão linear final. Além disso, os valores de

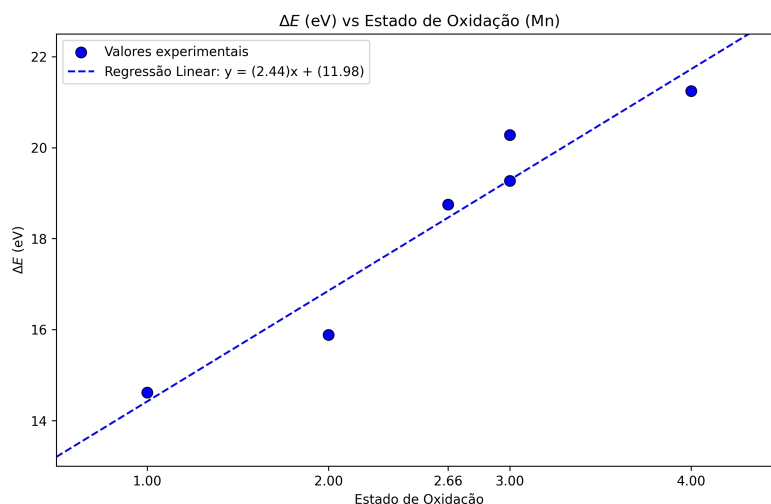


Figura 6: Regressão linear encontrada para o Manganês

NOX = 0 referente aos espectros metálicos resultaram em valores inesperados e provavelmente não eram condizentes com o método, por isso as regressões não os incluíram.

Essas dispersões podem ser interpretadas com base em diferença nas medidas e nos parâmetros experimentais como temperatura e solvente utilizado, porém além disso, mesmo podendo ter o mesmo estado de oxidação, o  $\Delta E$  apresentou variações dependendo de quais átomos estão ligados com o elemento analisado. Isso indica que o método é eficaz e pode trazer *insights* sobre o estado de oxidação, mas não é suficiente para descrever totalmente o ambiente químico do elemento em questão, ou seja, não é um *fingerprint*. Além disso, da maneira que o método foi implementado, a região de pré-pico é substituída por uma função mais uniforme por meio de uma interpolação, e por isso perdemos informações sobre esta região. O que pode ser uma hipótese do porquê os estados metálicos estão apresentando divergências com a regressão linear e os outros resultados.



## 4 Conclusão e Aprendizado

### 4.1 Implementação de método para determinação de máximos e mínimos em espectros XANES

A implementação do método Rising Sun foi bem sucedida, no entanto, alguns ajustes quanto a sua generalização devem ser realizados. Com as informações de máximos e mínimos, a possibilidade de comparação de espectros XAS a partir de métodos de aprendizado de máquina, que utilizem simulações e descritores, torna-se mais factível em futuras aplicações. Isso, com o principal objetivo de facilitar a comparação de espectros e auxiliar futuros usuários da linha QUATI na interpretação rápida e precisa de seus dados, aumentando a eficiência em estudos de materiais.

### 4.2 Implementação de método para determinação de estados de oxidação em espectros XANES

Por meio da implementação computacional do método de *Capehart*[1], foi possível gerar uma regressão linear para cada um dos compostos contendo elementos no intervalo de  $[21 \leq Z \leq 30 | Z \in \mathbb{N}]$  e perceber a relação linear de Wong [7], ou seja, conseguimos determinar os estados de oxidação de compostos 3d, revelando padrões de deslocamento ( $\Delta E$ ) que refletem as variações no número de oxidação dos elementos. As limitações observadas estão muito relacionadas com as maneiras experimentais que os arquivos *.xdi* foram medidos, e abrem margem para uma análise mais cuidadosa posterior, controlando esses parâmetros para curvas de calibração ainda mais precisas do que as encontradas no presente trabalho. As diferenças também podem ser relacionadas à influência de outros átomos ligados ao elemento analisado, que podem impactar o valor mesmo para o mesmo estado de oxidação. Apesar dessas limitações, o método provou ser útil para estudar o estado de oxidação, mas ainda necessita de aprimoramentos para fornecer uma descrição mais completa do ambiente químico e dos efeitos locais ao redor dos elementos analisados.

Com a combinação do método de determinação do número de oxidação

(NOX) e o método de mínimos e máximos utilizando o Rising Sun, abre-se a possibilidade de aplicar técnicas de machine learning para a identificação de elementos a partir de espectros XDI[9]. A obtenção precisa de características espectrais, como os valores de  $\Delta E$  e os pontos de mínimo e máximo, fornece um conjunto de dados robusto que pode ser utilizado para treinar modelos preditivos, permitindo a automação e a análise de grandes volumes de dados espectroscópicos com maior eficiência e precisão.

O trabalho foi muito importante e interessante para o desenvolvimento de habilidades computacionais, leitura/comentário de códigos e organização de arquivos. Além de, é claro, ter sido fundamental para a aplicação de conhecimentos de física e computação previamente conhecidos em um ambiente novo, e a aquisição de novos conhecimentos nessas áreas.

## 5 Agradecimentos

A Santiago Figueroa e James Almeida pela orientação, explicações, aprofundamentos e paciência. Ao Rafael Yoshimura, da QUATI, pela companhia às sextas-feiras. Ao Rafael Monteiro, um dos desenvolvedores do método Rising-Sun, por estar disposto auxiliar com o projeto tirando dúvidas sobre o método. Ao Diogo Carvalho que auxiliou com erros de código e análise de resultados. Ao João Otávio e Carlos Daniel que auxiliaram a criação do job submetido para o HPC, uma grande aventura. E ao Gustavo Uchôa e Izaque Oliveira, pelo trabalho de mineração de dados para alimentar a Cruzeiro do Sul Database e pelo apoio durante o projeto.

## Referências

- [1] T. W. Capehart et al. “X-ray-absorption edge shifts in rare-earth–transition-metal compounds”. Em: *Phys. Rev. B* 52 (11 set. de 1995), pp. 7907–7914. DOI: [10.1103/PhysRevB.52.7907](https://doi.org/10.1103/PhysRevB.52.7907). URL: <https://link.aps.org/doi/10.1103/PhysRevB.52.7907>.
- [2] Martina Fracchia et al. “Time-Resolved X-ray Absorption Spectroscopy in (Photo)Electrochemistry”. Em: *Surfaces* 1.1 (2018), pp. 138–150. ISSN: 2571-9637. DOI: [10.3390/surfaces1010011](https://doi.org/10.3390/surfaces1010011). URL: <https://www.mdpi.com/2571-9637/1/1/11>.

- [3] Matthew Newville. “Fundamentals of XAFS”. Em: *Reviews in Mineralogy and Geochemistry* 78.1 (jan. de 2014), pp. 33–74. ISSN: 1529-6466. DOI: [10.2138/rmg.2014.78.2](https://doi.org/10.2138/rmg.2014.78.2). eprint: [https://pubs.geoscienceworld.org/msa/rimg/article-pdf/78/1/33/2952366/33\\_REV078C02.pdf](https://pubs.geoscienceworld.org/msa/rimg/article-pdf/78/1/33/2952366/33_REV078C02.pdf). URL: <https://doi.org/10.2138/rmg.2014.78.2>.
- [4] James Almeida. *Cruzeiro do Sul Database*. <https://github.com/jamesmalmeida/Cruzeiro-do-Sul-Database>. A repository for X-ray Absorption Spectroscopy (XAS) datasets. 2024. URL: <https://github.com/jamesmalmeida/Cruzeiro-do-Sul-Database>.
- [5] X-ray Absorption Fine Structure (XAFS) community. *XDI File Format Specification*. <https://docs.xrayabsorption.org/xaslib/xdi.html>. Documentation for the X-ray Data Interchange (XDI) file format. 2024. URL: <https://docs.xrayabsorption.org/xaslib/xdi.html>.
- [6] Rafael Monteiro, Itsuki Miyazato e Keisuke Takahashi. “Rising Sun Envelope Method: An Automatic and Accurate Peak Location Technique for XANES Measurements”. Em: *The Journal of Physical Chemistry A* 124.9 (2020). PMID: 32013431, pp. 1754–1762. DOI: [10.1021/acs.jpca.9b11712](https://doi.org/10.1021/acs.jpca.9b11712). eprint: <https://doi.org/10.1021/acs.jpca.9b11712>. URL: <https://doi.org/10.1021/acs.jpca.9b11712>.
- [7] J. Wong et al. “K-edge absorption spectra of selected vanadium compounds”. Em: *Phys. Rev. B* 30 (10 nov. de 1984), pp. 5596–5610. DOI: [10.1103/PhysRevB.30.5596](https://doi.org/10.1103/PhysRevB.30.5596). URL: <https://link.aps.org/doi/10.1103/PhysRevB.30.5596>.
- [8] Augusto E. Mejía Gómez et al. “Crystal structure, cobalt and iron speciation and oxygen non-stoichiometry of La<sub>0.6</sub>Sr<sub>0.4</sub>Co<sub>1-y</sub>FeyO<sub>3</sub>- nanorods for IT-SOFC cathodes”. Em: *Journal of Alloys and Compounds* 817 (2020), p. 153250. ISSN: 0925-8388. DOI: <https://doi.org/10.1016/j.jallcom.2019.153250>. URL: <https://www.sciencedirect.com/science/article/pii/S0925838819344962>.
- [9] A. A. Guda et al. “Understanding X-ray absorption spectra by means of descriptors and machine learning algorithms”. Em: *npj Computational Materials* 7.1 (2021), p. 203. ISSN: 2057-3960. DOI: [10.1038/s41524-021-00664-9](https://doi.org/10.1038/s41524-021-00664-9). URL: <https://doi.org/10.1038/s41524-021-00664-9>.