

## FEW-GROUP CROSS SECTIONS LIBRARY BY ACTIVE LEARNING WITH SPLINE KERNELS

**Szames E., Ammar K., Tomatis D. and Martinez J.M.**

CEA, DEN, DM2S Service d'études des réacteurs et de mathématiques appliquées,  
 Université Paris-Saclay F-91191 Gif-sur-Yvette, France

esteban.szames@cea.fr, karim.ammar@cea.fr,  
 daniele.tomatis@cea.fr, jean-marc.martinez@cea.fr

### ABSTRACT

This work deals with the representation of homogenized few-groups cross sections libraries by machine learning. A Reproducing Kernel Hilbert Space (RKHS) is used for different Pool Active Learning strategies to obtain an optimal support. Specifically a spline kernel is used and results are compared to multi-linear interpolation as used in industry, discussing the reduction of the library size and of the overall performance. A standard PWR fuel assembly provides the use case (OECD-NEA Burn-up Credit Criticality Benchmark [1]).

KEYWORDS: Homogenized Cross Sections, Machine Learning, Kernel Methods, Pool Active Learning

### 1. INTRODUCTION

Few-group cross sections are obtained through an homogenization process from transport calculations, that compute the neutronic flux with a detailed discretization in energy and space. Cross sections are considered real valued scalar functions in a  $d$ -dimensional rectangular domain. Without loss of generality we normalize it into  $\mathcal{X} = [0, 1]^d$ . For every lattice calculation point, the cross section set  $\mathcal{Y} = \{\sigma(\bar{x}) \rightarrow \mathbb{R}, \bar{x} \in \mathcal{X}\}$  of size  $|\mathcal{Y}| = i \times r \times g$  is obtained for every reaction type  $r$ , group  $g$  and isotope  $i$ . These are generally smooth functions with possible strong variations in localized regions. Discrete samples  $\mathcal{X}_S = \{\bar{x}_i \in \mathcal{X}\}$  define the *support*  $\mathcal{S} = \{\sigma(\bar{x}_i), \bar{x}_i \in \mathcal{X}_S\}$ , which is the available information to build the model. These are called “data sites” or “learning space”. The modeling effort consists in finding the approximations  $\hat{\sigma}(\bar{x}|\mathcal{S}) \simeq \sigma, \forall \sigma \in \mathcal{Y}$ .

Macroscopic cross sections ( $\Sigma$ ) are the input data of core calculations and allow to calculate the infinite multiplication factor ( $k_\infty$ ). They are determined by a weighted sum of microscopic cross sections:

$$\Sigma_{r,g} = \Sigma_{r,g}^{res} + \sum_{i=1}^{N_I} \sigma_{i,r,g} C_i; \quad k_\infty = \frac{\nu_1 \Sigma_{f,1} (\Sigma_{a,2} + \Sigma_{2 \rightarrow 1}) / \Sigma_{1 \rightarrow 2} + \nu_2 \Sigma_{f,2}}{(\Sigma_{a,1} + \Sigma_{1 \rightarrow 2}) (\Sigma_{a,2} + \Sigma_{2 \rightarrow 1}) / \Sigma_{1 \rightarrow 2} - \Sigma_{2 \rightarrow 1}}, \quad (1)$$

being  $C_i$  the  $i$ -th isotope's concentration for  $N_I$  specialized isotopes. Ordinary reactor studies for fuel cycle optimization, transient simulations and core design need to resolve the multi-physics coupling with other computer codes where cross sections models have to deal with ever-growing volumes of data.

## 2. MODELING CHALLENGES AND STATE OF THE ART

In many industry implementations cross sections are modeled by a first order piece-wise polynomial interpolation here called Multi-Linear (ML) [2]. A Cartesian sampling rule for  $\mathcal{X}_S$  is normally used with support points *shared* by all cross sections. ML does not require an actual modeling phase, data are just stored, and can be fastly interpolated on demand. However, this advantage comes at a cost: the library of size  $|\mathcal{X}_S| \times |\mathcal{Y}|$  grows exponentially with increasing dimensions ( $|\mathcal{X}_S| \propto N^d$ ), which is known as the ‘‘Curse of Dimensionality’’ [3]. Higher order approximation spaces, projection into sub-libraries and regression with global polynomials [3] mitigate this to some extent still requiring a very dense support, with the possibility of either localized errors or high polynomial degrees for some cross sections [4]. Sparse Grid sampling rules, especially when anisotropic, have successfully reduced the library size though the design is left to the user [5]. Nested regular grids were needed, being actually a subset of a full grid, and Chebyshev sampling rules were used to deal with the Runge phenomenon resulting from global polynomials. Another example of the use of anisotropic sampling can be found in [6], where the model is optimized to increase the accuracy of  $k_\infty$  though limited to a Cartesian grid. In [7] Artificial Neural Networks were employed with high accuracy but at a significant computational cost. In [8] the ‘‘Empirical Interpolation Method’’ was used to obtain the support by reducing the model’s basis errors. Yet, the supervised learning implementation is severely intertwined with the expansion space: support points candidates result from a compositional procedure based on the error of the eigen-functions. The sampling is restrained to a ‘‘Tucker grid’’ used to solve the integral quadrature and subjected to a second selection process for lowering the amount of additional lattice calculations. A posteriori shrinking is then performed to reject unimportant terms as in many other techniques [3], [5]. The key ideas addressed in this work are:

- A full grid support suffers not only from the ‘‘Curse of Dimensionality’’, but includes also a significant amount of unnecessary data as evidenced by the low errors attained with sparse grids [5], by the variability in the amount of regression coefficients in [4] and by the use of a posteriori shrinking [8].
- Benefits of using unstructured support with respect to full grids have been suggested but not examined exhaustively, certainly not for a traditional function space that can be confronted to multi-linear [8].
- Though some effort has been made in studying the strategies to optimize the support’s anisotropy, no comparative analysis is to be found on selection criteria based on  $\sigma$ ,  $\Sigma$  and  $k_\infty$  and their resulting errors, nor the interplay between cross section sharing support and the supervised learning procedure.

## 3. KERNEL METHODS

Provided the support  $\mathcal{S}$  of size  $N = |\mathcal{S}|$ , our task is to find the approximation  $\hat{\sigma}$  to  $\sigma$  in some function space  $\mathcal{F}$

$$\hat{\sigma}(\bar{x}) = \arg \min_{\varphi \in \mathcal{F}} \frac{1}{N} \sum_{i=1}^N (\sigma(\bar{x}_i) - \varphi(\bar{x}_i))^2 + \lambda \|\varphi\|_{\mathcal{F}}^2, \text{ with } \lambda \in \mathbb{R}_{\geq 0}. \quad (2)$$

Kernel Methods [9] are a very general framework that allows to pose in a vector space, equipped with a norm, a vast quantity of machine learning problems. They are based on symmetric positive-

definite scalar function  $k : (x, y) \in \mathbb{R}^2 \rightarrow \mathbb{R}$  called *kernel* that generates a Reproducing Kernel Hilbert Space (RKHS)  $\mathcal{H}_k$ . Indeed with  $\varphi \in \mathcal{H}_k$  the solution of Eq. 2 is found by solving the linear system  $(K + \lambda I)\bar{\alpha} = \bar{\sigma}$ , with  $\bar{\sigma}_i = \sigma(\bar{x}_i)$  and the kernel matrix  $K_{i,j} = \mathcal{K}(\bar{x}_i, \bar{x}_j)$  of size  $N \times N$ , being  $\mathcal{K}(\bar{x}, \bar{y}) = \prod_{i=1}^d k(x_i, y_i)$ . The regression problem in Eq. 2 may be greatly simplified when expressed in terms of the kernel and this is known as the *kernel trick*. The approximation is then the linear combination of partially evaluated kernels

$$\hat{\sigma}(\bar{x}) = \sum_{i=1}^N \alpha_i \mathcal{K}(\bar{x}, \bar{x}_i), \quad (3)$$

as stated in the *Representer Theorem* [9]. Although  $\mathcal{H}_k \subset \mathcal{F}$ , quite exotic  $\mathcal{H}_k$  can be obtained enabling highly customized function spaces as exploited in [8]. Yet, we choose a first-order piecewise polynomial space (with the regularization term  $\lambda = 0$ ) reproduced by the *spline kernel* [10]:

$$k(x, y) = 1 + \max(x, y). \quad (4)$$

The two main drawbacks of Kernels Methods are the condition number of the matrix  $K$  and the possible large amount of terms in Eq. 3 for  $|\bar{\alpha}| = |\mathcal{X}_S|$ . A first order spline kernel reduces the conditioning problem. Though maybe suboptimal, this function space facilitates the comparison to multi-linear interpolation on a full grid and provides a bounded interpolation error. Higher order function spaces will be analyzed in a future paper.

#### 4. POOL ACTIVE LEARNING

Kernel Methods impose no condition on the sampling rules for the domain, thus enabling the use of Pool Active Learning (AL). Lattice calculations provide a *pool*  $\mathcal{S}_P = \{\sigma(x_i), x_i \in \mathcal{X}_P\}$  from which an optimal sampling  $\mathcal{X}_\dagger \subset \mathcal{X}_P$  can be extracted, allowing to define an optimal support  $\mathcal{S}_\dagger$ . This is done by computing the extrema of a *loss function*  $\mathcal{L}$  which lays at the heart of the AL process and has been merely treated implicitly in other works [11]. Two approximation scenarios are considered: with and without a shared support for the cross sections, as presented in pseudo code in Algorithms 1 and 2, respectively. In both cases the model starts with an initial support  $\mathcal{S}_0 \in \mathcal{S}_P$  chosen randomly and loss function values are computed within the loop to find new optimal points  $\bar{x}_\dagger \in \mathcal{X}_P$ . In the former case a maximum size  $|\mathcal{X}_\dagger|$  is selected, whereas a target error  $\delta_\sigma$  is specified per each cross section in the latter. The loss functions studied in this work, each generating a different model, are presented in Table 2d. We define the isotope's "importance" in view of their contribution to  $\Sigma$  as  $I_\sigma = I_{i,r,g} = C_i \sigma_{i,r,g} / \Sigma_{r,g}$ .

#### 5. RESULTS

The transport code APOLLO2.8 was used to generate cross section data for the OECD-NEA Burn-up Credit Criticality Benchmark in which material and geometrical specifications are fully available [1]. It consists of a typical PWR fuel assembly composed of  $17 \times 17$  UO<sub>2</sub> fuel rods with 4% w/o <sup>235</sup>U enrichment and with 25 guide tubes. We consider  $d = 3$  linearly independent variables: the Burnup (0 - 45 GWd/t), the Fuel Temperature (500 - 1000 °C) and the Boron Concentration (300 - 800 ppm) sampled in a  $170 \times 16 \times 16$  Cartesian grid. The specialized isotopes are  $\mathcal{I} = \{^{109}\text{Ag}, ^{243}\text{Am}, ^{153}\text{Eu}, ^{155}\text{Gd}, ^{95}\text{Mo}, ^{143}\text{Nd}, ^{145}\text{Nd}, ^{237}\text{Np}, ^{238}\text{Pu}, ^{239}\text{Pu}, ^{240}\text{Pu}, ^{241}\text{Pu}, ^{242}\text{Pu},$

**Data:**  $\mathcal{S}_p, \mathcal{S}_0, b, \mathcal{L}$

```

1  $\mathcal{X}_\dagger = \mathcal{X}_0$ 
2 while  $|\mathcal{X}_\dagger| \leq b$  do
3    $\hat{y} = \hat{y}(\bar{x}|\mathcal{S}_\dagger), \forall y \in \mathcal{Y}$ 
4    $\bar{x}_\dagger = \arg \max_{\bar{x} \in \mathcal{X}_p \setminus \mathcal{X}_\dagger} \|\mathcal{L}(\hat{y}, y)\|_\infty$ 
5    $\mathcal{X}_\dagger = \mathcal{X}_\dagger \cup \bar{x}_\dagger$ 
6 end
Result: An optimal  $\mathcal{X}_\dagger$  with respect to  $\mathcal{L}$  of size  $b$ .
```

**Algorithm 1:** Cross sections share support;  $y$  can be any  $\sigma, \Sigma$  or  $k_\infty$ .

**Data:**  $\mathcal{S}_p, \mathcal{S}_0, \delta_\sigma$

```

1 for  $y \in \mathcal{Y}$  do
2    $\mathcal{X}_{\sigma,\dagger} = \mathcal{X}_0$ 
3   while  $err \leq \delta_\sigma$  do
4      $\hat{\sigma} = \hat{\sigma}(\bar{x}|\mathcal{S}_{\sigma,\dagger})$ 
5      $err = \|\Delta\sigma(\bar{x})\|_\infty$ 
6      $\bar{x}_{\sigma,\dagger} = \arg \max_{\bar{x} \in \mathcal{X}_p \setminus \mathcal{X}_{\sigma,\dagger}} err$ 
7      $\mathcal{X}_{\sigma,\dagger} = \mathcal{X}_{\sigma,\dagger} \cup \bar{x}_{\sigma,\dagger}$ 
8   end
9 end
Result: An optimal set  $\mathcal{X}_{\sigma,\dagger}$ .
```

**Algorithm 2:** Independent support optimization for every  $\sigma$ .

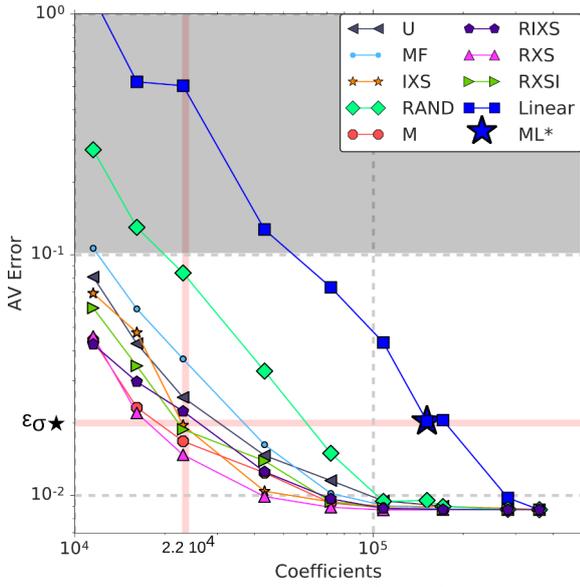
**Figure 1: Pseudo code for Pool Active Learning. The support  $\mathcal{S}$  includes the sample  $\mathcal{X}$  on which it is defined.**

$^{103}\text{Rh}, ^{101}\text{Ru}, ^{147}\text{Sm}, ^{149}\text{Sm}, ^{150}\text{Sm}, ^{151}\text{Sm}, ^{152}\text{Sm}, ^{99}\text{Tc}, ^{234}\text{U}, ^{235}\text{U}, ^{236}\text{U}, ^{238}\text{U}, ^{133}\text{Cs}, ^{135}\text{Xe}, ^{135}\text{I}$  for the reactions  $\nu\sigma_{f,1}, \nu\sigma_{f,2}, \sigma_{a,1}, \sigma_{a,2}, \sigma_{1 \rightarrow 2}, \sigma_{2 \rightarrow 1}$  resulting in  $|\mathcal{Y}| = 128$  (fission products have no fission reaction). Lattice calculations data are divided into two disjoint sets the pool  $\mathcal{S}_p$  and the test  $\mathcal{T}$  ( $\mathcal{S}_p \cap \mathcal{T} = \emptyset$ ). Test points properly cover the domain and are fixed independently of the cross section's support resulting in a stable error evaluation. The microscopic cross section relative error is defined as  $\Delta\sigma = 100(\hat{\sigma}/\sigma - 1)$ . An arithmetic average of the error over the test set is  $\epsilon_\sigma = \sum_{i=1}^{|\mathcal{T}|} abs(\Delta\sigma_i)/|\mathcal{T}|$ , and over all cross sections is  $\varepsilon_\sigma = \sum_{\sigma=1}^{|\mathcal{Y}|} abs(\epsilon_\sigma)/|\mathcal{Y}|$ . Modern cross section target errors are between  $10^{-2}$  and  $10^{-1}$  in % [5]. Similar error definitions are used for  $\Sigma$  and  $k_\infty$ . With any representation of the form in Eq. 3, the library size amounts to the number of coefficients to store. For a cross section approximated by Kernel Methods with AL,  $|\bar{\alpha}| = |\mathcal{X}_\dagger|$  and if they share support  $\sum_{\sigma=1}^{|\mathcal{Y}|} |\bar{\alpha}_\sigma| = |\mathcal{X}_\dagger| \times |\mathcal{Y}|$ .

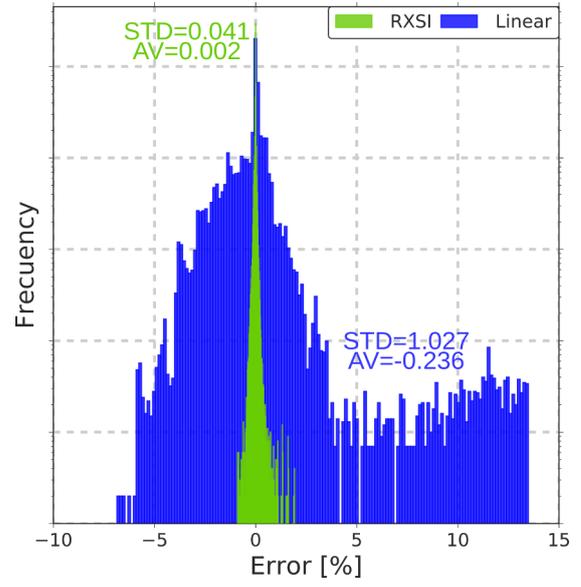
### 5.1. Active learning with shared support

The trend of  $\epsilon_\sigma$  with increasing library size is presented in Fig 2a. ML always uses a full grid and  $\mathcal{S}_*$  marks a special discretization based on ‘‘industry’s expert knowledge’’ for this type of fuel assembly. This defines a comparison line of constant error  $\varepsilon_{\sigma,*}$  while the other points show the global tendencies. Modern error limits are delimited in gray [5]. The errors decrease monotonically bounded by  $\sigma(x)$  and as  $\mathcal{X}_S \rightarrow \mathcal{X}_p$ , they converge since the same first order approximation space is used for all the models. It can be noticed that the error does not reach zero, being the *useful* information of  $\mathcal{S}_p$  limited and  $\mathcal{T}$  fixed. For comparing different  $\mathcal{L}$  a library size of  $2.2 \times 10^4$  coefficients is big enough to appreciate the effect of the AL with errors not yet leveled by enforcing the majority of  $\mathcal{X}_p$ . We start by noting that a random selection (RAND) already produces a significant improvement as expected: collinearity undermines the informativeness of an approximation’s support [4]. In a way this isolates the effect of using an unstructured support which reduces the library size by half. All AL strategies further decrease it as presented in Fig. 3b.

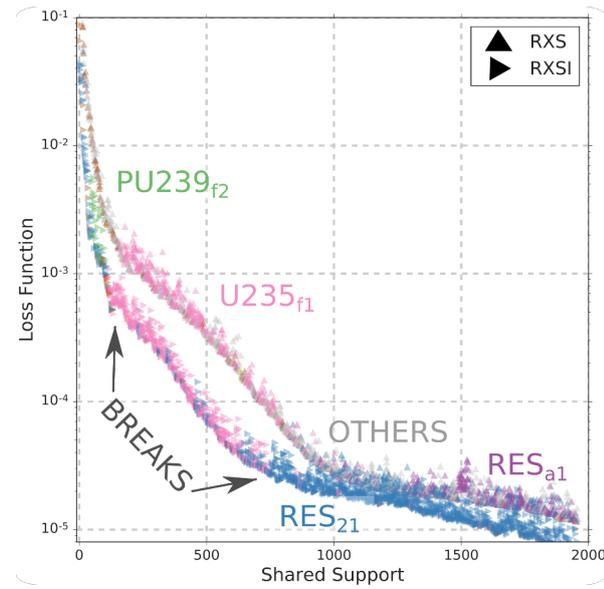
The lowest  $\sigma$  error is obtained when considering the relative errors for the whole set (RXS). Ad-



(a) Average  $\sigma$  error with shared support.



(b) Error histogram at  $2.2 \times 10^4$  coefficients.



(c) Cross section participating in the AL.

Name	Loss Function	Description	$\mathcal{I}_\sigma$	Errors
RAND	RAND	Random selection.	-	-
RXS	$\arg \max_{\sigma \in \mathcal{Y}} \Delta \sigma / \sigma$	All cross sections.	No	Rel.
RXSI	$\arg \max_{\sigma \in \mathcal{Y}} I_\sigma \Delta \sigma / \sigma$	Important cross sections.	Yes	Rel.
IXS	$\arg \max_{\sigma \in \mathcal{Y}} I_\sigma \Delta \sigma$	All cross sections.	Yes	Abs.
U	$ \hat{U}_{a,2}^{235} - U_{a,2}^{235} $	Single cross section.	No	Abs.
RIXS	$\arg \max_{\sigma \in \mathcal{Y} \text{ if } \max(I_\sigma) > 0.2} I_\sigma \Delta \sigma / \sigma$	Restrained to cross sections relevant to $\Sigma$	Yes	Rel.
M	$\arg \max_{\Sigma \in \mathcal{Y}_\Sigma} \Delta \Sigma$	All macroscopic cross sections.	-	Abs.
MF	$ \hat{k}_\infty - k_\infty $	Multiplication factor.	-	Abs.

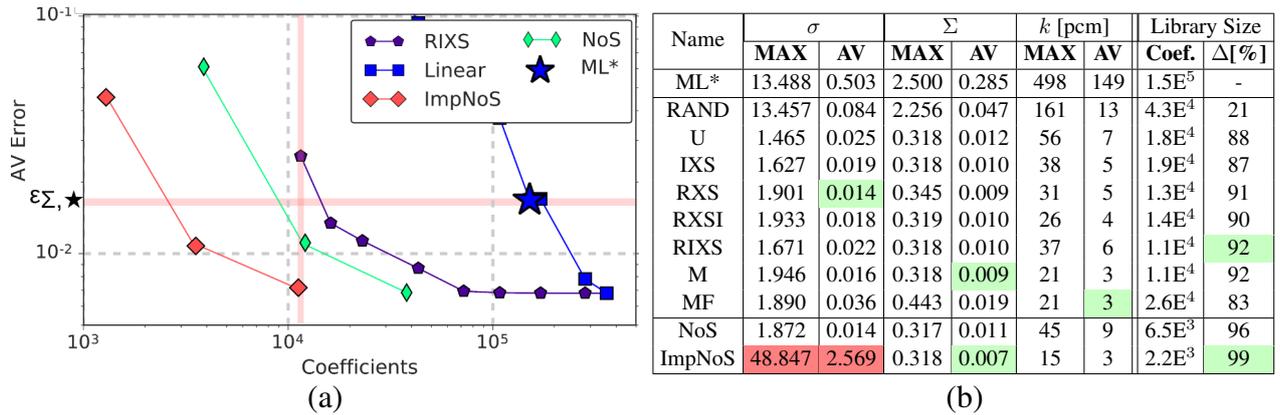
(d)  $\mathcal{L}(y, \hat{y})$  with  $\sigma$ ,  $\Sigma$  or  $k_\infty$  as  $y$ .

**Figure 2: Cross section modeling with Kernel Methods and Active Learning.**

ditionally weighting with the importance (RXSI) does not improve it. Its histogram is shown in Fig. 2b characterized by centered means, normally distributed shape, small standard deviations and no error trails, unlike ML for that library size. Similar histograms are noticed for the other AL strategies. If absolute errors are considered, even if weighted by the importance (IXS), we find that the only cross section participating in the AL is  $^{135}\text{Xe}_{a2}$  since its absolute values are significantly bigger than the others. In this case the AL effort is entirely wasted with  $\hat{\sigma}$  computed  $|\mathcal{Y}|$  times. Still, the error profile differs little from the others methods, in the same way as optimizing for single cross section (U). This shows that the active learning process is quite robust with respect

to loss function selection. To better understand this behavior in Fig.2c loss function values are plotted showing the cross section participating in the selection of each  $\bar{x}_\dagger$  for RXS and RXSI. The use of  $I_\sigma$  reduces significantly their amount: while RXS considers about 50 cross sections (marked in gray and being 1/3 of  $|\mathcal{Y}|$ ) RXSI only  $\sim 10$ .  $\mathcal{L}$  values are bigger for RXS than for RXSI (since  $I_\sigma \leq 1$  and  $I_\sigma \sim 0$  for a large amount of cross section) but the overall profile and error stagnation ( $|\mathcal{X}_S| > 1000$ ) remains the same, suggesting that the useful information is being extracted at a similar rate from  $\mathcal{X}_P$ . The profiles present clear breaks in the derivative (at a support of 200 and 1000) that can be understood as a change in the “relevance” of the points being added to the model and could even be used as a stop criterion for Algo 1. We further notice a clustering effect where the cross sections participating (independently) in the AL, end up segregated in a few blocks mainly dominated by  $U_{f1}^{235}$  and  $RES_{2 \rightarrow 1}$ . These exhibit jumps in  $\mathcal{L}$  due to the ongoing AL process. Being the support shared,  $\bar{x}_\dagger$  is added to all cross sections inducing plateaus in their loss function profiles (not shown here). The optimization effort is thus subordinated to a very small cross section set while forcing the model to incorporate large amounts of unnecessary data as explained in Section 2. Finally the importance can be used to preselect a small cross section subset (RIXS) to performed AL obtaining similar results at a lower computational cost.

A  $\mathcal{L}$  based on macroscopic cross sections (M) exhibits a particular good performance for  $\Sigma$  and one based in the multiplication factor (MF) in  $k_\infty$  as can be seen in Fig. 3b. Though AL could be considered an “off-line” task, performed only once during cross section preparation, we note that IXS, RXS, RXSI, M and MF require the computation of the whole cross section set unlike RIXS or U that achieve a similar error.



**Figure 3: (a) Average  $\Sigma$  error without shared support. (b) Errors for AL models for  $\sigma$ ,  $\Sigma$ ,  $k_\infty$ . Errors at constant support  $\sim 1 \times 10^4$  coefficients. Library size at constant error  $\epsilon_{\Sigma,*}$ .**

### 5.2. Active learning without shared support

Since only the coefficients need to be stored we consider Algo. 2 enabling an independent AL process for each cross section. We consider two “No Shared Support” methods defined by  $\delta_\sigma = \delta$  (NoS) and by weighting with the importances  $\delta_\sigma = \delta I_\sigma$  (ImpNoS) with  $\delta = 0.01, 0.001, 0.0001$  whose  $\epsilon_\Sigma$  are presented in Fig. 3a. For  $\epsilon_{\Sigma,*}$  the library size is further reduced up to another order of magnitude arriving at 1% of ML in a Cartesian grid. However for ImpNoS we notice high  $\epsilon_\sigma$  since the AL only focuses in cross sections relevant to  $\Sigma$ . This is explained by Fig 4b where the amount of coefficients relative to the library size is marked by the area for every cross section. Only  $^{235}U_{a1}$ ,  $^{235}U_{f1}$ ,  $RES_{1 \rightarrow 2}$  and  $RES_{2 \rightarrow 1}$  surpass 300 coefficients, thus achieving low errors, while this is the norm for the majority of cross sections in NoS presented in Fig. 4a.



ior at a smaller computational cost. Indeed by using unstructured supports  $|\mathcal{X}_S| \propto N^d$  effectively dealing with the “Curse of dimensionality”. We point out the need of considering relative errors to have meaningful comparisons. In the current Python-F2PY-Fortran implementation with vectorized kernel evaluations and sharing the kernel evaluation matrix an evaluation time of  $T_\sigma \sim 10^{-5}$  seconds per cross section was observed. Further code optimization may achieve higher computational performance. A shared support penalizes the AL as evidenced by the low participation rate (clustering) of cross sections where many are forced to incorporate large volumes of unnecessary data. By dropping this constraint,  $\sum_\sigma |\bar{\alpha}_\sigma| \neq |\mathcal{X}_\dagger| \times |\mathcal{Y}|$  and the AL process is able to exploit the different cross section modeling requirements. Weighting with  $I_\sigma$  resulted in a satisfying library composition, where the coefficient’s cardinality intuitively follows important cross sections though high errors were noted in many cross sections negligible to  $\Sigma$ . Our methodology for cross section representation not sharing the support allowed a reduction of up to two orders in magnitude, i.e. up to 1% of the original library size, for target accuracies as used in industry.

## REFERENCES

- [1] A. Barreau. “Burn-up Credit Criticality Benchmark.” Technical Report Phase, IID, OECD-NEA (2006).
- [2] U. Grundmann et al. “DYN3D version 3.2 - code for calculation of transients in light water reactors (LWR) with hexagonal or quadratic fuel elements - description of models and methods -.” *Wissenschaftlich-Technische Berichte* (2019).
- [3] R. Zivanovic and P. Bokov. “Cross-section parameterization of the pebble bed modular reactor using the dimension-wise expansion model.” *Annals of Nuclear Energy*, **volume 37**, pp. 1763–1773 (2010).
- [4] J. Dufek. “Building the nodal nuclear data dependences in a many-dimensional state-variable space.” *Annals of Nuclear Energy*, **volume 38**(7), pp. 1569 – 1577 (2011).
- [5] D. Botes and P. M. Bokov. “Polynomial interpolation of few-group neutron cross sections on sparse grids.” *Annals of Nuclear Energy*, **volume 64**, pp. 156 – 168 (2014).
- [6] C. Sanchez et al. “Optimization of multidimensional cross-section tables for few-group core calculations.” *Annals of Nuclear Energy*, **volume 69**, pp. 226 – 237 (2014).
- [7] E. Szames, K. Ammar, D. Tomatis, and J. Martinez. “Few-group cross sections library compression by artificial neural networks.” *Submitted to proceeding of the conference: Physics of reactor conference, Cambridge, United Kingdom* (2019).
- [8] H. L. Thi et al. “Use cases of Tucker decomposition method for reconstruction of neutron macroscopic cross-sections.” *Annals of Nuclear Energy*, **volume 109**, pp. 284 – 297 (2017).
- [9] B. Scholkopf and A. J. Smola. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press, Cambridge, MA, USA (2001).
- [10] G. Wahba. *Spline Models for Observational Data*. Society for Industrial and Applied Mathematics, Philadelphia (1990).
- [11] D. Wu. “Pool-Based Sequential Active Learning for Regression.” *IEEE transactions on neural networks and learning systems* (2018).