Reconstruction of few-group homogenized cross section by kernel method and active learning

Ea. Szames, D. Tomatis, K. Ammar, Jm. Martinez

RECONSTRUCTION OF FEW-GROUP HOMOGENIZED CROSS SECTIONS BY KERNEL METHODS AND ACTIVE LEARNING

Esteban Szames, Daniele Tomatis, Karim Ammar and Jean-Marc Martinez
DEN, Service d'études des réacteurs et de mathématiques appliquées
CEA, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France
esteban.szames@cea.fr, daniele.tomatis@cea.fr, karim.ammar@cea.fr, jean-marc.martinez@cea.fr

ABSTRACT

This work deals with the approximation of homogenized few-groups cross sections by kernel methods. Different kernels types are used in conjunction with pool active learning to optimize the cross section's support. They are compared to multi-linear interpolation and multi-variate splines, similar to industry [1]. A standard PWR fuel assembly is provided by the OECD-NEA Burn-up Credit Criticality Benchmark, Phase-IID, to evaluate their performances [2].

KEYWORDS: Few-Group Cross Sections, Kernel Methods, Active Learning, Spline Interpolation

1. INTRODUCTION

Core calculations are performed on spatially coarse meshes and the variety of reactor states are reproduced by a few-group cross section model [3]. Homogenized cross sections are obtained from transport calculations with detailed energy and geometry discretizations which are not feasible to perform on-the-fly in the core simulator. Consequently they are calculated a priori and stored in libraries for reconstruction on demand. This is known as the two-step calculation approach. Usually a selected amount of specialized isotopes I are tracked by the core and macroscopic cross sections are reconstructed as

$$\Sigma_{r,g} = \Sigma^{res}_{r,g} + \sum_{i=1}^{I} \sigma_{i,r,g} C_i, \tag{1}$$

for every reaction r and group g being C_i the i-isotope's concentration. These are the most important fissile and absorbers while the rest are lumped together in a residual cross section $\Sigma^{res}_{r,g}$. As common in LWR we use a two-energy group structure with a cutoff of 0.625 eV.

2. FEW-GROUP CROSS SECTIONS MODELS

Cross sections are considered as real valued scalar functions evaluated in a d-dimensional rectangle domain $\mathcal{D}$ encompassing the wide range of possible reactor conditions. The general functionalization is $\sigma_{i,r,g}(\boldsymbol{x})$ with $\boldsymbol{x} = (x_1, \ldots, x_d) \in \mathcal{D}$ of linearly independent variables called "state-parameters": physical quantities showing high influence on the reaction rates. In this work we consider typical choices for PWRs, such as the burnup (0-4.5 GWd/t), fuel temperature (500-1000 °C) and boron concentration (300-800 ppm), hence $d = 3$. Without loss of generality this domain is mapped into a unit hyper cube $D = [0, 1]^d$.

Cross sections approximation consists on finding $\hat{\sigma}_{\mathcal{T}} : D \to \mathbb{R}$ for generally smooth functions exhibiting locality in relatively small parts of the domain (e.g. BOC) and a low order dependence among the

variables. The approximation's support is noted as $\mathcal{T} = \{[\sigma(\boldsymbol{\tau}_1), \boldsymbol{\tau}_1], \ldots, [\sigma(\boldsymbol{\tau}_n), \boldsymbol{\tau}_n]\}$ of size n obtained from sampling the state-parameter space with transport calculations. We denote as $\boldsymbol{\tau}$ a point $\boldsymbol{x} \in D$ used as support. Depending on the research field $\mathcal{T}$ may be called "support space", "data sites", "response vectors" or "learning space".

Multi-linear interpolation in full grids have been commonly used in cross section representation [4]. Though reconstruction is simple and fast, large library sizes may be needed to reach the desired accuracy [5]. Higher order approximations such as Splines [6], usually combined with projection into dedicated sub-libraries, mitigate this to some extent by needing a smaller support. To circumvent a Cartesian construction of the domain sparse grid using Legendre polynomials with Chebyshev nodes have been used allowing some degree of anisotropy in the sampling rules [7] as to profit from the changing cross section's complexity with the state-parameters. Recently, machine learning techniques have been investigated to aid the reduction of the support size. Optimization by considering the infinite multiplication factor's error has been performed though restrained to a regular (Cartesian) grid [8]. Unstructured supports have been explored as well but only in view of the microscopic cross section's error [9].

3. POOL ACTIVE LEARNING

The transport data forms a *pool* $\mathcal{P}$ from which we extract the most informative support $\mathcal{T}^\dagger \subset \mathcal{P}$ for the learner, i.e. for $\hat{\sigma}_\mathcal{T}$. This optimal data set is found by computing the extrema of an objective function f in a process called pool active learning (AL) [10]. The function f lays at the heart of the AL process and we consider a reduction in weighted microscopic cross section errors as presented in Algo. 1 using

$$w_{i,r,g} = \frac{C_i \sigma_{i,r,g}}{\Sigma_{r,g}}, \tag{2}$$

that assess each isotope's contribution to $\Sigma_{r,g}$ being $w_{i,r,g} \leq 1$. The set of such weights is noted as $\mathcal{W}$. A starter support $\mathcal{T}_0$ and a budget b that fixes the cardinality of $\mathcal{T}^\dagger$ must be provided as well. Then, in a sequential procedure, $\hat{\sigma}_\mathcal{T}$ is computed enabling to select the best support point $\boldsymbol{\tau}^\dagger \in \mathcal{P}$ with respect to f in each iteration. An optimal *unstructured* (or scattered) domain is thus found, without any user intervention or expert knowledge.

4. KERNEL METHODS

Kernel methods are frequently used in machine learning problems such as regression or classification where they are known as Support Vector Machines [11]. As the name suggests they are based on scalar kernel function $k : (x, y) \in \mathbb{R}^2$ which are symmetric and non-negative defining a Hilbert function space $\mathcal{H}_k$ called the Reproducing Kernel Hilbert Space (RKHS). The approximation $\hat{\sigma}_\mathcal{T}$ in $\mathcal{H}_k$ results from minimizing

$$\hat{\sigma}_\mathcal{T} = \underset{\varphi \in \mathcal{H}_k}{\arg\min} \frac{1}{n} \sum_{i=1}^{n} \left(\sigma(\boldsymbol{\tau}_i) - \varphi(\boldsymbol{\tau}_i)\right)^2 + \lambda \|\varphi\|^2_{\mathcal{H}_k}, \tag{3}$$

where the regularization term is defined by the norm in $\mathcal{H}_k$ and $\lambda \geq 0$. The *kernel trick* allows to evaluate $\hat{\sigma}_\mathcal{T}$ very easily by solving the linear system $(K + \lambda n I)\boldsymbol{\alpha} = \boldsymbol{\sigma}$, with the kernel matrix $K_{i,j} = k(\boldsymbol{\tau}_i, \boldsymbol{\tau}_j)$ and $\boldsymbol{\sigma}_i = \sigma(\boldsymbol{\tau}_i)$, $1 \leq i, j \leq n$. Then $\hat{\sigma}_\mathcal{T}(\boldsymbol{x}) = \sum_{i=1}^{n} \alpha_i k(\boldsymbol{x}, \boldsymbol{\tau}_i)$ with $k(\boldsymbol{x}, \boldsymbol{y}) = \prod_{j=1}^{d} k_j(x_j, y_j)$. We utilize for each dimension the kernel

$$k_j(x_j, y_j) = \sum_{l=0}^{m} \frac{1}{(l!)^2} B_l(x_j) B_l(y_j) + \frac{(-1)^{(m+1)}}{(2m)!} B_{2m}(|x_j - y_j|), \tag{4}$$

where B_l is the l-th Bernoulli polynomial and $1 \leq j \leq d$. A central feature of the RKHS is that the corresponding space $\mathcal{H}_K$ is composed of functions in the Sobolev space that are m times differentiable, similar to the smoothing splines [12].

Data: $\mathcal{P}$, $\mathcal{W}$, $\mathcal{T}_0$, b
initialize: $\hat{\sigma}_{\mathcal{T}_0}$;
while $\#\mathcal{T} \leq b$ **do**
  compute: $f = w(\hat{\sigma}_{\mathcal{T}} - \sigma), \forall, i, r, g$;
  find: $\boldsymbol{\tau}^\dagger = \arg\max_{\boldsymbol{x} \in \mathcal{P} \setminus \mathcal{T}} ||f||_\infty$;
  extend: $\mathcal{T} = \mathcal{T} \cup \boldsymbol{\tau}^\dagger$;
  compute: $\hat{\sigma}_{\mathcal{T}}$;
end
Result: An optimal $\mathcal{T}^\dagger \subset \mathcal{P}$ with respecto to $\mathcal{W}$ for the budget b that defines $\hat{\sigma}_{\mathcal{T}^\dagger}$.

Algorithm 1: Pool active learning algorithm.

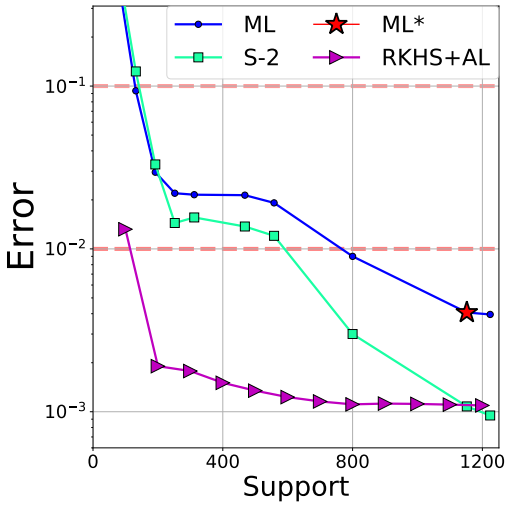

Figure 1. Macroscopic error [%] for multi-linear (ML), quadratic spline (S-2) and RKHS with active learning.

5. VALIDATION ON A TOPICAL CASE

Lastly, we also consider more classical expansion techniques, such as the linear combination of the tensor product of univariate basis functions

$$\hat{\sigma}_{\mathcal{T}}(\boldsymbol{x}) = \sum_{\boldsymbol{j}} \alpha_{\boldsymbol{j}} \phi_{\boldsymbol{j}}(\boldsymbol{x}) = \sum_{j_1=1}^{N_1} \cdots \sum_{j_d=1}^{N_d} \alpha_{j_1,\ldots,j_d} \prod_{i=1}^{d} \psi_{i,j_i}(x_i), \tag{5}$$

where we use B-splines as basis for $\phi(\boldsymbol{x}) = \prod_{i=1}^{d} \psi_i(x_i)$ in a full grid [13]. The coefficients $\alpha_{\boldsymbol{j}}$ are obtained by imposing the interpolation condition $\hat{\sigma}_{\mathcal{T}}(\boldsymbol{\tau}) = \sigma(\boldsymbol{\tau}), \forall \boldsymbol{\tau} \in \mathcal{T}$ for each cross section. Multi-linear interpolation can be considered as a Spline of degree 1.

A standard PWR fuel assembly of 17 $\times$ 17 fuel pins was used with the transport code APOLLO2 [14] to assess the quality of the representation methods. Critical flux calculations were performed using 281-group cross section library based on JEFF-3.1 with a B1 leakage model. In Fig. 1 the arithmetic average of the relative macroscopic cross section error $\epsilon_\Sigma = 100|\hat{\Sigma} - \Sigma|/\Sigma$ is presented for the fission, absorption and scattering cross sections in D.

Multi-linear (ML), quadratic Spline (S-2) and the RKHS (with $m = 2$) curves are plotted in function of the support's size. Max. and average target errors are shown in dotted lines. A red star marks the typical discretization used for these type of fuel assemblies with multi-linear interpolation in full grids. Relative to the Spline, Multi-linear errors are bigger being a first order approximation. The use of AL greatly accelerates the error's convergence rate for RKHS when compared to a full grid discretization reducing the required support by a significant amount. Eventually when the totality of the pool is included $\mathcal{T}^\dagger \sim \mathcal{P}$ and the RKHS error overlaps with S-2, being both 2$^{\text{nd}}$ order methods.

6. DISCUSSION

A discussion of the results obtained from the approximation methods and their applicability, scalability with the number of dimensions and limitations will be address here. A thorough exploration of the error's composition will be addressed in this section with particular attention to the effect of a Σ-driven support construction in σ representations. Other kernels will be presented discussing the benefits and drawbacks in regards to library size and evaluation speed. It will also be address the choice of the norm in AL, e.g. $||w\Delta\sigma||_1$, $||w\Delta\sigma||_2$, etc. exploring their effect in the selection of points. Some comments about the implementation of the methods will be address as well.

REFERENCES

[1] G. Hobson et al. "ARTEMISTM core simulator: latest developments." In *International Conference on Supercomputing in Nuclear Applications Monte Carlo* (2014).

[2] A. Barreau. "Burn-up Credit Criticality Benchmark." Technical Report Phase, IID, OECD-NEA (2006).

[3] K. Ivanov and M. Avramova. "Challenges in coupled thermalhydraulics and neutronics simulations for LWR safety analysis." *Annals of Nuclear Energy*, **volume 34**, pp. 501–513 (2007).

[4] J. K. Watson and K. N. Ivanov. "Improved cross-section modeling methodology for coupled three-dimensional transient simulations." *Annals of Nuclear Energy*, **volume 29**(8), pp. 937 – 966 (2002).

[5] J. Dufek. "Building the nodal nuclear data dependences in a many-dimensional state-variable space." (2011).

[6] J. G. Herriot and C. H. Reinsch. "Algorithm 600: Translation of Algorithm 507. Procedures for Quintic Natural Spline Interpolation." *ACM Trans Math Softw*, **volume 9**, pp. 258–259 (1983).

[7] D. Botes and P. M. Bokov. "Polynomial interpolation of few-group neutron cross sections on sparse grids." (2014).

[8] S. Snchez-Cervera, N. Garca-Herranz, J. Herrero, and O. Cabellos. "Optimization of multidimensional cross-section tables for few-group core calculations." *Annals of Nuclear Energy*, **volume 69**, pp. 226 – 237 (2014).

[9] T. H. Luu, M. Guillo, Y. Maday, and P. Gurin. "Use cases of Tucker decomposition method for reconstruction of neutron macroscopic cross-sections." *Annals of Nuclear Energy*, **volume 109**, pp. 284 – 297 (2017).

[10] D. Wu. "Pool-Based Sequential Active Learning for Regression." *IEEE transactions on neural networks and learning systems* (2018).

[11] B. Scholkopf and A. J. Smola. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press, Cambridge, MA, USA (2001).

[12] G. Wahba. *Spline Models for Observational Data*. Society for Industrial and Applied Mathematics, Philadelphia (1990).

[13] C. de Boor. "A practical guide to splines." (2001).

[14] R. S. at al. "APOLLO2 Year 2010." *Nucl Eng and Technology*, **volume 42**, pp. 474–499 (2010).