Driving Global Inference for New Physics with Machine Learning, Big Data and Large-scale Statistical Simulation

Matthias Danninger (Simon Fraser University),
Jonathan Cornell (Weber State University),
Nazila Mahmoudi (CERN),
Gregory Martinez (University of California, Los Angeles).

August 1 – August 6 2022

1 Overview of the Field

The Standard Model of particle physics and the ΛCDM model of cosmology [1] encapsulate our knowledge of the constituents of the universe, and how they caused the universe to evolve to its present state. Despite their incredible joint success in explaining experimental data, both theories suffer from a number of experimental and theoretical challenges that continue to steer the direction of global research. For example, the Standard Model famously does not include gravity, cannot explain the matter-antimatter asymmetry in our universe, cannot naturally explain the fact that the Higgs boson mass is much lower than the Planck scale and cannot provide an explanation for the dark matter in our universe that is a key component of the ΛCDM model. The latter model, meanwhile, suffers from discrepancies in the inferred value of the Hubble constant H_0 by CMB and supernovae data [16], clustering σ_8 [11] by CMB and weak lensing data and curvature Ω_k [9, 8] by CMB and lensing BAO data (in addition to discrepancies between CMB datasets [10]). To uncover the laws of physics that ultimately resolve these issues, we need to combine insights from a wide range of previous, current and future experiments, in fields as disparate as collider, neutrino and flavour physics, cosmology and astrophysics. This can be accomplished within the framework of Bayesian or frequentist global statistical fits, which can be used to quantify how well a given physics model fits known data, determine which ranges of the model parameters are preferred, and quantify tensions between divergent datasets.

First released in 2017, the open-source Global and Modular Beyond-the-Standard-Model Inference Tool (GAMBIT) [2] performs global statistical fits of generic theories of new physics – so-called Beyond-the-Standard-Model (BSM) theories – including those that can resolve the cosmological and particle physics anomalies and tensions. GAMBIT is capable of producing results in both the Bayesian and frequentist statistical frameworks, and is easily extendible to new physics models and new experimental datasets. A fully modular design enables much of the code to be reused when changing the theoretical model of interest, and there are separate modules for handling collider observables, dark matter direct and indirect search data, cosmological measurements, flavour physics observables, precision measurements of Standard Model processes and neutrino observables. Additional theoretical modules enable the mass and decay spectra of new particles to be calculated, including an automated approach for dealing with running masses and couplings. GAMBIT includes a wide variety of efficient sampling algorithms for posterior evaluation and optimisation, and ensures computational efficiency through massive, multi-level parallelisation, both of the sampling algorithms and individual likelihood calculations.

2 Recent Developments and Open Problems

GAMBIT has been used to study a wide range of physical theories, including supersymmetric scenarios, various effective theories of WIMP dark matter (including contact interaction-based EFTs and simplified models with explicit mediators), models with extended neutrino sectors and axion theories. To build on this success, it is not sufficient to simply keep pace with experimental developments and pursue ever wider classes of theory. Rather, we need to advance the state of the art in sampling technology, use machine-learning to speed up simulations and/or change the way that inference is performed (e.g. using ”likelihood-free inference” techniques), and develop myriad new approaches to make global fits in high dimensional parameter spaces tractable.

3 Presentation Highlights and Scientific Progress

The purpose of our workshop was to bring GAMBIT experts together with machine learning experts and statisticians to tackle the above listed open problems, and to plan the next year of activities for the GAMBIT community. In particular, we discussed and outlined the structural code developments required in each part of the framework to successfully develop and include machine learning and new sampling methods into GAMBIT. Below we list a few particular highlights in this development that was compiled by workshop participants leading the respective discussions during the BIRS workshop.

3.1 ScannerBit

by Will Handley, Anders Kvellestad and Gregory Martinez

Data-driven exploration of computationally expensive physics models requires 1) adaptive parameter sampling, 2) efficient parallelisation, and 3) the application of fast approximations for expensive computations where possible. These were all major topics during the workshop. In the GAMBIT Community, the Scanner working group, together with the Core working group (see below), are in charge of the development of new approaches to these challenges, and their implementation in the ScannerBit module of the GAMBIT code framework. During the workshop, members of these working groups led sessions of presentations and discussions on the above challenges.

Below we highlight three topics of recent developments in the Scanner working group related closely to challenges 1) and 3) above, pointing out the progress made during the workshop and the open questions that were identified.

3.1.1 Fast-slow scanners

A long-term goal of GAMBIT since at least GAMBIT IX in Germany has been to properly integrate “fast-slow” scanning into the code. Such scanning can occur when a likelihood has a hierarchy of speeds, in that holding certain “slow” parameters constant, and only changing “faster” parameters yields a much faster likelihood re-computation. Such a facility is only possible if the likelihood is capable of cacheing portions of the calculation between calls, and if those computationally expensive portions depend on only a (small) subset of the parameters. This is exploited to great effect in cosmology codes such as CAMB and CLASS, where the cosmological parameters governing the properties of the universe are “slow” (requiring recomputation of the transfer functions from the lengthy Boltzmann simulations of the universe from start to today), whilst the nuisance parameters associated with galactic foregrounds or satellite calibration are “fast”, since these only affect the final step in the computation of the likelihoods. Codes such as COSMOMC, cobaya, cosmosis and

MontePython all make use of this fast-slow hierarchy, enabling their Bayesian scanners to scale to much higher dimensions than would otherwise be possible.

GAMBIT as a codebase is in a unique position, in that its organised and modularised structure is ideally suited in principle to automatically cache any or all portions of the likelihood. Fast-slow hierarchies have only been used in Bayesian scanners, but Frequentist or differential scanners potentially could benefit even more, allowing near-immediate optimisation of any fast parameters at any call.

To implement fast-slow scanning within GAMBIT requires work on two-fronts. First within the Core, to create a caching and timing infrastructure, and second within Scanners to exploit this. Currently only one scanner (PolyChord) natively supports fast-slow, but if the facility were made available this would very soon change.

We had a lengthy whiteboard discussion between Core, Cosmo & ScannerBit members on the correct abstraction from the Core side, and the desired API from the scanner side, as well as some discussions on how other scanners might exploit this ability. This was uniquely facilitate by an in-person meeting with a whiteboard an many members of both teams present. Further work has since been spurred in this regard, and we hope to accomplish a preliminary fast-slow scanning structure by the time of the next GAMBIT meeting.

3.1.2 Continual learning

Regression based on machine learning (ML) can be used to tackle an expensive calculation or simulation by training a fast approximation, or emulator, to replace the heavy computation. However, some computations depend intimately on frequently updated input data, as is the for instance with simulations of LHC searches for new particles. In such cases, pre-trained emulators will have very limited reusability. Also, since an adaptive parameter sampler will narrow in on the parameter subspace in best agreement with current data, attempting to pre-train precise global emulators for such data-dependent simulations will represent a significant waste of computational resources. An alternative approach is to both train and apply an emulator on the fly. This is an example of continual learning (CL), an approach to ML mostly explored within robotics and autonomous systems.

A long-term goal for the Scanner working group is to establish CL as a key technique for lowering the computational expense of, and extending the useful outputs of, GAMBIT parameter scans. By mimicking a "man-in-the-middle attack" we can apply CL during parameter scans, independent of the choice of parameter sampling algorithm. The CL algorithm operates as a middle layer, intercepting the communication between the scanner, which picks new parameter points θ, and the evaluation of the expensive target function $f(\theta)$.

By learning from the stream of corresponding inputs θ and outputs $f(\theta)$, the CL algorithm can gradually train a fast emulator for the target function in the local regions of θ-space explored by the scanner. When the CL algorithm intercepts a point θ for which its estimate $\hat{f}(\theta)$ has sufficiently low uncertainty, it can short-circuit the expensive function evaluation by passing $\hat{f}$ back to the scanner.

For this approach to be useful, the CL algorithm must satisfy the following desiderata: i) fast training, ii) fast evaluation, iii) reliable per-point uncertainty estimates, and, iv) limited need for hand-tuning of the algorithm settings. These requirements can all be satisfied by a recent CL approach called Dividing Local Gaussian Processes (DLGP) [13, 12]. Standard Gaussian Processes (GP) regression already satisfies desiderata iii) and iv) above. However, the time to train a regular GP scales as n^3 with the number of training points n. This makes standard GP regression unusable for continual learning, which requires repeated training based on a 'never-ending' stream of training points. The DLGP approach circumvents this by using the stream of training data to dynamically grow a binary tree where each leaf is a small-size GP, each responsible for approximating the target function across a some parameter subspace, and each with a size cap $n \sim \mathcal{O}(100)$. This ensures that training and evaluation times are kept low, since only a few, small GPs must be evaluated and/or retrained for each new point. GAMBIT members Anders Kvellestad and Are Raklev are currently working with statisticians Riccardo De Bin and Timo Lohrmann on a project to develop a modified DLGP algorithm tailored especially for BSM studies and GAMBIT use. At the workshop, recent progress from this project was reported (see Fig. 1).

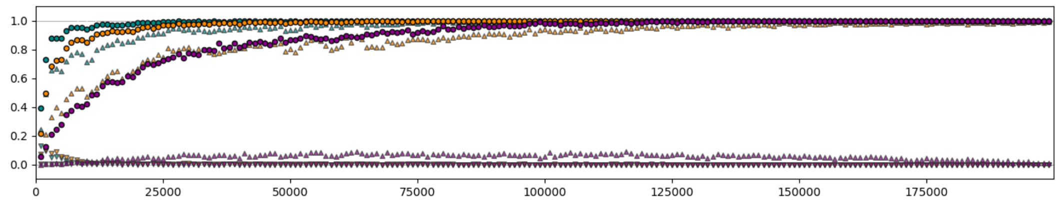

Figure 1: A preliminary test of the application of DLGP-based continual learning during a parameter scan collecting 200,000 samples in total. A six-dimensional Rosenbrock function was used as the target function for the DLGP algorithm. The scan was performed using a differential evolution algorithm. For each sampled parameter point the DLGP emulator was first queried for a fast prediction of the target function, and then afterwards trained on the true target function value. For each batch of 1000 samples, the circles show the fraction of those samples for which the DLGP prediction is within 10% (green), 5% (orange) and 1% (purple) of the true target function value. After $\sim 80,000$ samples the DLGP prediction is within 5% of the true value for more than 99% of subsequently sampled points, and after $\sim 120,000$ samples the prediction has a less than 1% error for 99% of all new samples. Considering that a differential evolution-based GAMBIT study of a six-parameter BSM theory typically involves target function evaluations at millions of parameter points, this test shows that the application of continual learning has the potential for providing significant computational gains.

At the workshop we started planning and developing code for how to interface this new CL algorithm to ScannerBit. A key requirement for this interface is that it should be sufficiently general and modular so as to allow plug-and-play interfacing also of future algorithms for continual learning, similar to how ScannerBit currently is interfaced to a range of different algorithms for scanning parameter spaces. An illustration of the planned user interface for the GAMBIT CL system can be seen in Fig. 2.

A central feature of CL system planned during the workshop will be the ability to store emulators trained during one GAMBIT scan, so that they can be reused, and further refined, during subsequent GAMBIT runs of the BSM theory. Since a large-scale BSM global fit with GAMBIT typically requires a campaign of 10–20 runs, such a CL system can help ensure that the computational resources of later runs are not wasted on performing expensive computations in parts of parameter space already explored in high detail by the earlier scans. A further extension discussed during the workshop is a system for automatically wrapping trained CL emulators in a small stand-alone interface. This will enable the final, trained CL emulators to easily be made publicly available as part of the data release for the given global fit study, to enable computationally cheap follow-up studies.

The computational challenge of exploring high-dimensional and computationally expensive new physics theories necessitates the use of high-performance computing (HPC) systems. GAMBIT studies are regularly run on large HPC systems, with single GAMBIT runs using up towards 10,000 cores. A currently open question is how the future CL functionality of GAMBIT best can make use of this massive parallelisation. Several different possible approaches were discussed during the workshop, e.g. combining independently trained DLGP emulators for the same target function, e.g. via a boosting, random forest or the Generalised Robust Bayesian Committee Machine approach [14], and learning emulators for many different target functions in parallel. Exploration and comparison of different approaches to massively parallelised continual learning will form an important topic for future research.

In order to speed up the convergence of a global fit, all that is needed is a reliable CL emulator for the single, total likelihood function of the given parameter scan. However, in order to understand the physics that explains why particular parameter regions are preferred by data, we need to know the function values for a large number of the physics predictions and/or likelihood components that

```yaml
ObsLikes:

  - capability: some_slow_computation
    purpose:    LogLike
    - options:
        train_emulator: true
        apply_emulator: true
        emulator_threshold: 0.01

  - capability: another_slow_computation
    purpose:    Observable
    - options:
        train_emulator: true
        apply_emulator: false

  - capability: yet_another_slow_computation
    purpose:    Observable
    - options:
        apply_emulator: true
        emulator_threshold: 0.05
```

Figure 2: An example of the planned user-interface for the GAMBIT continual learning system. From the YAML file that configures a GAMBIT run, the user will be able to specify for which computations emulators should be trained and/or used, and set the uncertainty threshold that determines when an emulator prediction is regarded sufficiently accurate to replace the corresponding expensive computation.

enter the computation of the total likelihood. For this reason, the ability to utilise the massive parallelisation of GAMBIT runs to train emulators for a large collection of different target functions simultaneously during a global fit will be a particularly important development. An illustration of why this is important can be seen in Fig. 3, which shows preliminary results from the most computationally expensive GAMBIT project so far. The project, which was the topic of detailed discussions during the workshop, considers the impact of LHC results on a particular class of supersymmetry scenarios. It will be the most comprehensive BSM global fit to date in terms of the detailed treatment of LHC physics: for every parameter point sampled in the scan we simulate 16 million LHC proton-proton collisions and pass them through emulations of the data filtering pipelines for a large number of ATLAS and CMS searches for new particles, and a large collection of ATLAS and CMS measurements of known Standard Model processes. The four panels in Fig. 3 show how the results of four different classes of such LHC analyses contribute to the total likelihood function. The ability to study such decompositions of the total likelihood function is crucial to understand a global fit result, and to use it to guide the design of future data analyses and experiments. Further possible extensions of DLGP and related continual learning methods were also discussed at the workshop, including

- possible frequentist approaches to emulator calibration [17];
- extending DLGP to multi-output GPs, useful for learning correlated functions [5], e.g. predictions for multiple signal regions in a collider search;
- dynamic updating of the GP kernel structures and the cap parameter; and
- approaches to “continual forgetting”, to achieve algorithms with a constant memory footprint.

3.1.3 Python scanners

by Gregory Martinez

In addition, the Scanner Working Group discussed the needs and requirements for future scanners in GAMBIT, with emphasis on the programming language used to develop GAMBIT scanners. Unfortunately, researchers and students typically are not well-versed in C/C++. This, in addition to added syntax required by the ScannerBit plugins, makes the learning curve for adding additional scanners prohibitively steep. Because of this, we decided that a Python plugin will need to be developed. We believe that this will entice further scanner development as most researchers and students are already familiar with the Python programming language, Python is simpler and easier

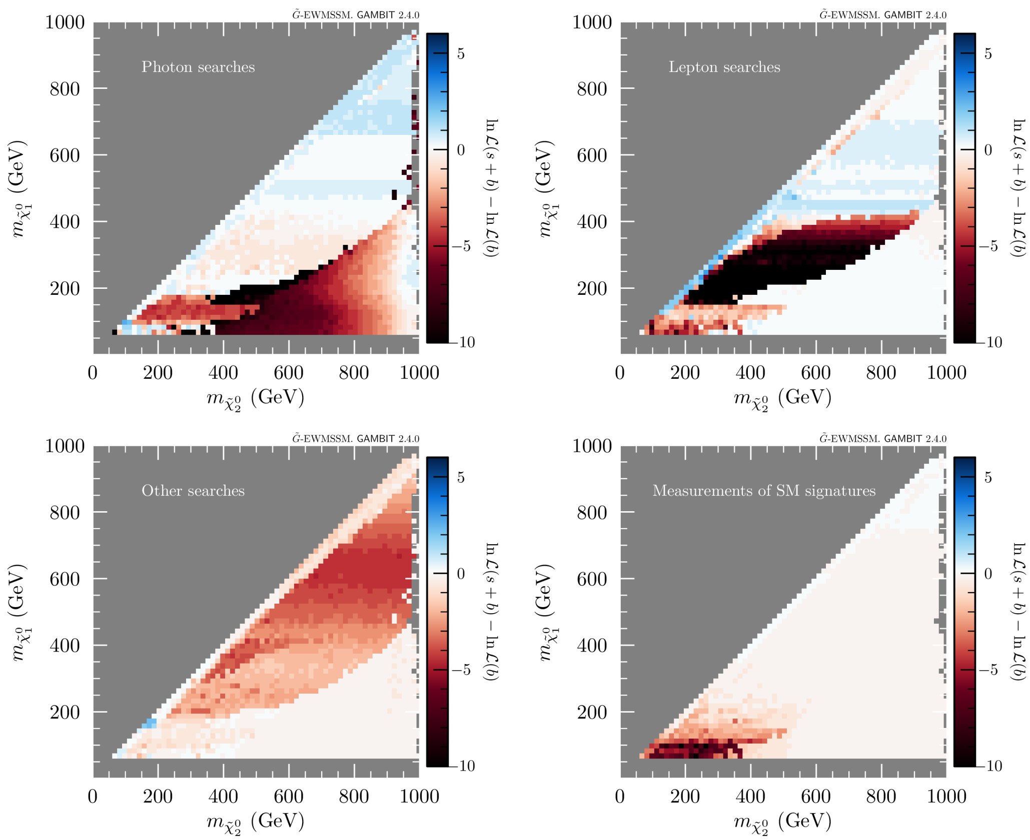

Figure 3: Preliminary results from an ongoing GAMBIT study of supersymmetry scenarios where the new particles within the energy reach of the LHC are four neutralinos, two charginos and a light gravitino. The four panels show the log-likelihood contributions from four classes of LHC searches and measurements, plotted for the overall highest-likelihood parameter sample within each bin of mass plane for the two lightest neutralinos.

to learn than C/C++, and there is already a huge ecosystem of packages written for Python that can be used for scanners.

Thus we decided that the Python plugin will need to have the following requirements:

- The Python plugin will need to allow Python scanners to be written completely in Python – e.g. no additional C/C++ code will be needed to be written in order to implement a new Python scanner.
- Python scanners must have the same inifile interface as other GAMBIT scanners so that users can easily use newly-implemented Python scanners.
- Python scanners must not interfere with already implemented GAMBIT Python backends.
- The Python plugin interface should be simple and intuitive, preferably similar to the C/C++ plugin interface already implemented as this will ease the transition to Python for users already familiar with the C/C++ plugin interface.

Additionally, development of the Python plugin began during the meeting where a proof-of-concept Python plugin was made that could start the Python interpreter, import a Python module, and run a simple Python function.

3.2 Core

by Tomas Gonzalo

The Core working group of the GAMBIT collaboration is responsible for the development and maintenance of the fundamental components of the GAMBIT framework. This includes, but it is not limited to, the build system, the backends, logs, printers, the diagnostics system, the dependency resolution and parser of YAML files, as well as the newly developed automatic code generation tool, GUM.

The most recent updates from the Core working group were consolidated into a new minor version release, GAMBIT 2.2. The released contained multiple physics updates, including support for cosmological axion-like particles, along with new and updated backends needed for their study. On the core side, the latest release most notably included the full support for Clang and AppleClang compilers, significantly simplifying the useage of GAMBIT on OSX systems. Various updates to BOSS and GUM also form part of this release, most of which were enhancements to support a wider variety of models and use cases, and some were minor bug fixes.

There were four major Core projects discussed during the meeting. The most advanced of which is the development of Template BOSS, a task undertaken by Tomas Gonzalo and Anders Kvellestad, with the help of Pawsey interns Zelun Li and Joel West. The bulk of this development is completed and the only remaining task is confirm that it works as expected with its principal use case, FlexibleSUSY. Another, Core exclusive, major project is the improvement of the HDF5 printers, which has been an occasional issue in recent times, and some progress has been done by Pascal Pelahi, from Pawsey. The redesign of SpecBit, a joint project with the Models/Precision working group, has been on hiatus for some time whilst Template BOSS was being developed, but can continue now that it is supposedly complete. The remaining tasks have been identified by Tomas Gonzalo and Peter Athron and will be undertaken in the near future. Lastly, the implementation of fast-slow scanners, a joint project with the Scanners working group was discussed in a dedicated session and individual tasks have been determined to push the progress.

Finally, all the Core developments, and some physics module developments, since the first release of GAMBIT, are being collated in a new Core paper, dubbed the GAMBIT 2 paper. Much remains to be documented, and other parts must wait until developments are finished, but the goal is clear and it is slowly taking shape.

3.3 ColliderBit

by Are Raklev

The ColliderBit working group is responsible for the consistent calculation of likelihoods from published collider searches for new physics, for example from the Large Hadron Collider(LHC), and for likelihoods connected to the measurement of the Higgs boson's parameters, as well as searches for new Higgs bosons.

Recent major updates from the ColliderBit working group includes the ability to use published full-likelihood models provided by the ATLAS experiment for some searches. The GAMBIT interface to the Contur and Rivet codes have also been improved, and the interface to HepMC updated. Issues with signal uncertainty and the accuracy of the likelihood profiling used inside of ColliderBit have been fixed.

During the meeting ongoing work on the changes needed to update the Monte Carlo event simulation to the new Pythia 8.3 code was discussed, and progressed to near completion. We have also discussed the use of multiple jet-collections needed for some LHC analysis, and the basic framework for this is in place in the HEPUtils framework, however, not yet connected to the analysis side. Work on this will continue over the summer, with a goal to use this in a planned paper focusing on the investigation of leptoquark models. Some risk to the speed of ColliderBit simulations were identified here. Because the running time of jet-algorithms is currently the time-limiting factor, the run time scales linearly with the number of requested jet collections.

Future efforts will focus on the ColliderBit Solo code to make the results of the ColliderBit group more useful to researchers outside of the GAMBIT Community. Focus will also be on speeding up the event generation, in particular connected to jet-clustering, PDF lookup, and phase space sampling. During the meeting we also had an extensive discussion on the potential of continual learning techniques to speed up signal predictions and the calculation of analysis likelihoods.

We plan to extend the capabilities of ColliderBit to handle collider searches for new long-lived particles (including vertex information), constraints from beam dump experiments, and to make the systems for interpolated yields and efficiencies more general and streamlined. To simplify future work more testing will be automated and cutflow checks will be made threadsafe. We will further investigate the possibility to use the MadGraph code internally in ColliderBit as an alternative to generating hard scattering processes at the LHC.

Several improvements on the collider cross section estimation system is also foreseen. Technical work will be undertaken to make the way we get initial cross section maximum estimates from Pythia more robust, we will simplify the structure of the LEP cross-section capability, and we will finalise interfaces to multiple codes that calculate higher-order cross sections: xsec, prospino, and salami.

3.4 DarkBit and CosmoBit

by Aaron Vincent & Will Handley

The purpose of the DarkBit module is to perform relevant calculations for dark matter-related observables: relic abundance computations are done by interfacing with the DarkSUSY and micrOMEGAs backends. Direct detection likelihoods are computed via DDCalc, and currently likelihoods are included for PICO-60, CRESST-II, CRESSST-III, DarkSide 50, LUX, PandaX and XENON1T. Constraints on indirect signals of dark matter are included via Fermi-LAT gamma-ray likelihoods, and the capture and annihilation of dark matter in the Sun can be computed by Capt'n General, with a corresponding neutrino signal at IceCube computed via DarkSusy. CosmoBit is a relatively recent addition [15], allowing for cosmological observables to be computed and tested, including the number of relativistic degrees of freedom in the Early Universe, alternate cosmic ionization histories, and various theories of inflation. Due to the natural overlap, DarkBit and CosmoBit development mainly takes place within a joint dark matter and cosmology working group. Recent work making use of DarkBit and CosmoBit include explorations of s-channel-mediated scalar and fermionic dark matter models [6], discussed at the workshop, and [4].

Ongoing work includes progress on a global fit of annual modulation experiments, with the joint goals of incorporating likelihoods from the many new modulation experiments (COSINE, SABRE, ANAIS), adding and expanding on features in DDCalc (effective operators, non-standard dark matter halo components), and determining whethere there is any place at all for a new physics explanation of DAMA to be hiding in the vast BSM parameter space. A second ongoing focus is on Sub-GeV dark matter models, which aim to explore features beyond those expected in the standard WIMP scenario.

The addition of neural networks to speed up likelihood evaluations was also discussed. The test case here is the computation of antiproton fluxes for arbitrary dark matter models, trained on results from the GALPROP cosmic ray propagation code. The interface between DarkRayNet and GAMBIT currently works, with a speedup from hours to seconds for marginalization over cosmic ray propagation parameters. Ongoing questions include how best to parallelize the process.

During the BIRS meeting, we established a suite of projects for further work. This has been encapsulated in (what we believe to be) a highly competitive DiRAC 15$^{\text{th}}$ Resource Allocation call for an April 2023 start. This is the UK's premier high-performance computing provision, and if succesful would further establish GAMBIT within the UK research funding structure. This would not have been possible without the initial discussion which took place at the BIRS meeting.

3.5 NeutrinoBit

by Chien Lin and Tomas Gonzalo

The goal of the NeutrinoBit working group is to perform global analyses of models with massive neutrinos. Recently the focus has been on the implementation of likelihood functions for neutrino oscillation measurements from various experiments, either from long-baseline accelerator experiments, reactors, solar and atmospheric neutrino oscillation experiments. Table 1 summaries the experiments currently included in GAMBIT. Many of the likelihood computations for these experiments have been brought to a reasonable accuracy, validation figures can be seen in Figure 4, while some are still being worked on or waiting to be improved. The development of the remaining experiments is delayed at the moment by several factors, for example, the limited access to experimental data and the highly experiment-dependency of detector and neutrino interaction models.

Long-baseline accelerator	Reactor	Solar	Atmospheric
T2K, NOvA, MINOS	Daya Bay, KamLAND	SNO	Super-K, IceCube

Table 1: The neutrino oscillation experiments currently included in NeutrinoBit.

In addition to the lack of data from various experiments, the computation speed of the calculations is another area of concern. Due to the complex nature of neutrino detector and the difficulties in modelling neutrino interactions, numerous systematic parameters related to the energy reconstruction and neutrino flux have to be introduced in the calculation. To maximise the likelihood, their corresponding parameter spaces have to be explored carefully which slows down the computation drastically. Several alleviating methods are currently being investigated, including the identification and exclusion of the weakest impacting nuisance parameters.

In addition to the study of neutrino oscillations, the Neutrino working group aims to expand to the study to other massive neutrino models. Several extensions of the neutrino oscillation study are already on the table, such as the inclusion of sterile neutrinos and non-standard interactions, and will be undertaken as soon as the minimal oscillation study is complete. The study of these extended models will be extremely useful on the goal to understand the various anomalies found in the oscillation of neutrinos on experiments such as MiniBoone and LSND, as well as for projected experiments such as DUNE and ORCA.

In parallel to the study of neutrino oscillations, the Neutrino working group is also involved in the study of models with heavy neutral leptons. These do not affect significantly the oscillation probability of the light neutrinos, but can be searched for at colliders or with precision experiments. A study of one such a model with three right-handed neutrinos has already been performed by the Neutrino working group [7], and a follow up to include updated constraints as well as predictions for future experiments is already underway.

3.6 SpecBit, DecayBit and PrecisionBit

by Peter Athron and Eliel Camargo-Molina

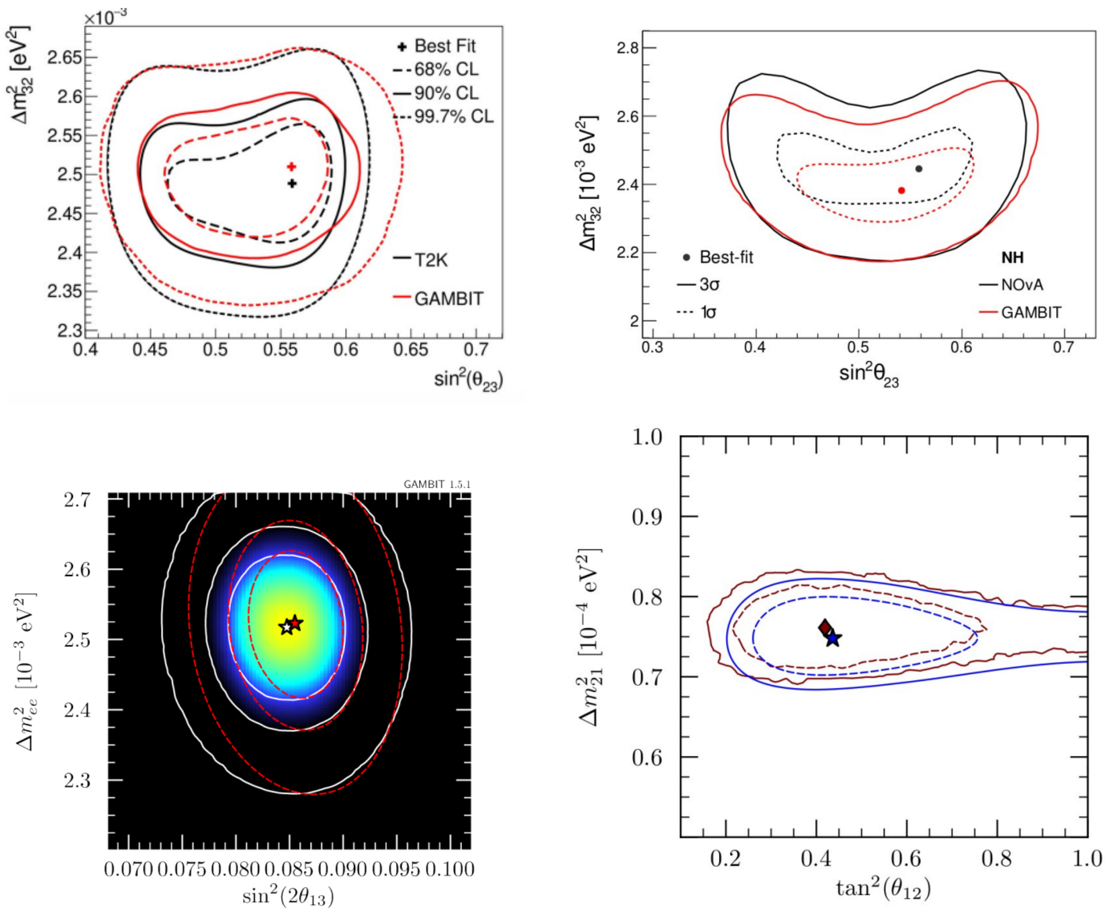

Figure 4: Validation of the likelihood functions for the T2K, NOvA, DayaBay and KamLAND experiments.

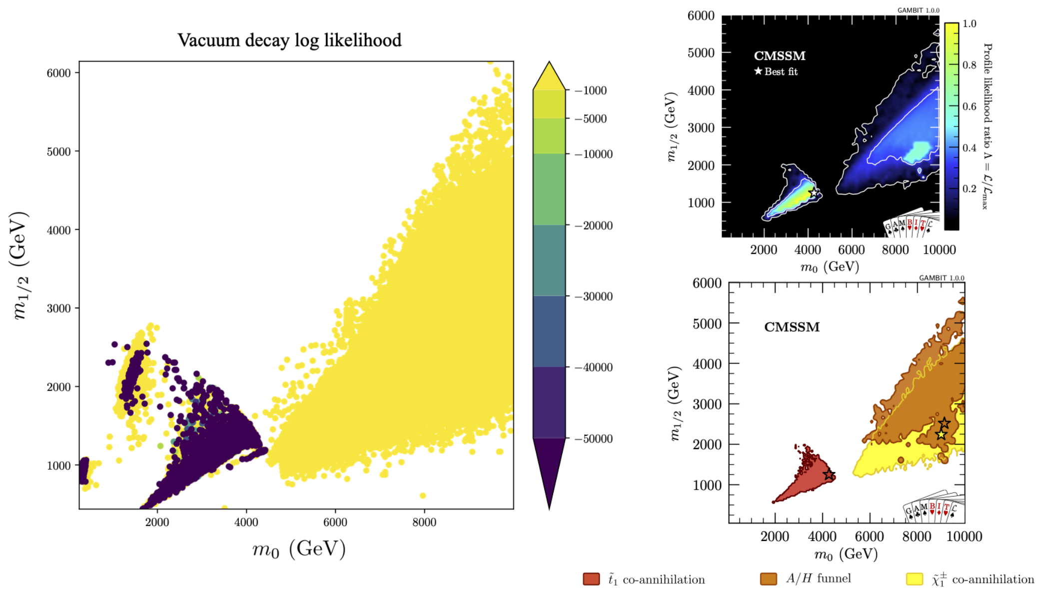

Figure 5: Preliminary comparison of global fit performed in [3] for the MSSM (right) and the log likelihood from considering vacuum decay to color- and charge- breaking minima calculated for the same points using Vevacious and GAMBIT (left). The region affected by vacuum decay largely overlaps with the region close to the best fit point.

The Precision Working Group is responsible for precision calculations that are not directly related to flavour, collider or dark matter. This includes calculation of the mass spectra, couplings and decays that are used in most of the other physics dedicated modules, as well as additional low energy precision calculations such as muon g-2 and electric dipole moments and precision calculations involving vacuum decay. We are responsible for SpecBit, DecayBit and PrecisionBit the latter of which we use to name the working group since all activities are connected to precision calculations.

This working group has been going through substantial changes. We are no longer responsible for the Models database, which is now under the responsibilities of the Core Working Group, leading to the switching the name from the Models Working group which emphases the key focus on our activities: precision predictions. Along with this the group is going through a period of revitalisation with existing members Peter Athron, Eliel Camargo-Molina, Tomas Gonzalo and Anders Kvellestad being joined by new people interested in our activities such as Douglas Jacob, Cristian Sierra, Roberto Ruiz, Yongcheng Wu, Martin White and Wei Su.

Key responsibilities of the group involve maintaining the existing calculations and updating backends when needed. In this respect we have an ongoing effort in the redesign of SpecBit, and recently have been updating GM2Calc to the latest version which can also do the two-Higgs doublet model, and updating FeynHiggs so we can make use of the state of the art on-shell MW calculation as well as the latest improvements in their Higgs mass calculation. The status of these was reported, fresh ideas were generated and some progress was made with FeynHiggs updating.

Progress with the new addition Vevacious was also discussed. For any theory and parameter values given by the user, Vevacious finds all the minima of the potential energy, adds thermal and quantum corrections, and calculates the lifetime of the vacuum reproducing the physics we observe. If the lifetime comes up much shorter than the age of the Universe, then the input parameters are excluded. This is a very challenging calculation which is often neglected in phenomenology due to the computational cost in terms of running the code (or any potential alternative) where people often instead resort to basic rules of thumb which are neither necessary nor sufficient conditions. As the calculations are computationally expensive and substantial runs have been done already for specific models by GAMBIT, machine learning solutions to this challenge ere discussed at the meeting. In particular, we discussed the case of the MSSM where we have a large dataset from previous work that shows that considering vacuum stability is crucial in identifying valid regions of the theory parameter space that (as shown in figure 5).

Significant new activities for the group were proposed and discussed in the meeting. Expanding our activities into proper global EW fits was proposed. We weighed up the benefits of making use of existing public EW global fit codes against developing our own in house set up. Another proposal is a timely investigation into the MW predictions of the MSSM following the recent CDF measurement, which is the most precise to date but has a significant deviation with the SM prediction and the previous measurements. We also discussed a long time plan and interest of members of the working group in phase transitions and its connections to gravitational waves and electroweak baryogenesis. Many ideas of physics projects to drive development of this were proposed and discussed.

4 Outcome of the Meeting

We made substantial progress on the common project of understanding which sections of the GAMBIT code can benefit and developed further to include fast ML techniques, aiding our physics analyses goals.

- We evaluated existing computational and statistical techniques used in global fits
- We discussed novel machine learning tools and techniques that could be suitable to be used for global fits
- We selected new computational tools and identified specific areas within GAMBIT for implementation
- Continual learning approaches will be further studied and developed
- We defined a clear publication plan for the next year and discussed for each project the crucial tasks that are required until publication. We started to draft some of these publications
- We made substantial progress on the planning for the paper describing the updated core software structure of GAMBIT since its initial release.
- We celebrated 10 years of successful collaboration and made new connections with new members (predominantly in the Canadian community).
- The opportunity to bring together 20 members of the global fit community in an in-person workshop after 2 years of COVID challenges was invaluable. This will strengthen the community in the next years to come.

References

[1] N. Aghanim et al. Planck 2018 results. VI. Cosmological parameters. Astron. Astrophys., 641:A6, 2020. [Erratum: Astron.Astrophys. 652, C4 (2021)].

[2] Peter Athron, Csaba Balazs, Torsten Bringmann, Andy Buckley, Marcin Chrząszcz, Jan Conrad, Jonathan M. Cornell, Lars A. Dal, Hugh Dickinson, Joakim Edsjö, Ben Farmer, Tomás E. Gonzalo, Paul Jackson, Abram Krislock, Anders Kvellestad, Johan Lundberg, James McKay, Farvah Mahmoudi, Gregory D. Martinez, Antje Putze, Are Raklev, Joachim Ripken, Christopher Rogan, Aldo Saavedra, Christopher Savage, Pat Scott, Seon-Hee Seo, Nicola Serra, Christoph Weniger, Martin White, and Sebastian Wild. GAMBIT: the global and modular beyond-the-standard-model inference tool. European Physical Journal C, 77(11):784, November 2017.

[3] Peter Athron et al. Global fits of GUT-scale SUSY models with GAMBIT. Eur. Phys. J. C, 77(12):824, 2017.

[4] Csaba Balázs et al. Cosmological constraints on decaying axion-like particles: a global analysis. 5 2022.

[5] Wessel Bruinsma, Eric Perim, William Tebbutt, Scott Hosking, Arno Solin, and Richard Turner. Scalable exact inference in multi-output Gaussian processes. In Proceedings of the 37th International Conference on Machine Learning, Proceedings of Machine Learning Research, pages 1190–1201. PMLR, 2020.

[6] Christopher Chang, Pat Scott, Tomás E. Gonzalo, Felix Kahlhoefer, Anders Kvellestad, and Martin White. Global fits of simplified models for dark matter with GAMBIT I. Scalar and fermionic models with s-channel vector mediators. 9 2022.

[7] Marcin Chrzaszcz, Marco Drewes, Tomás E. Gonzalo, Julia Harz, Suraj Krishnamurthy, and Christoph Weniger. A frequentist analysis of three right-handed neutrinos with GAMBIT. Eur. Phys. J. C, 80(6):569, 2020.

[8] Eleonora Di Valentino, Alessandro Melchiorri, and Joseph Silk. Planck evidence for a closed Universe and a possible crisis for cosmology. Nature Astron., 4(2):196–203, 2019.

[9] Will Handley. Curvature tension: evidence for a closed universe. Phys. Rev. D, 103(4):L041301, 2021.

[10] Will Handley and Pablo Lemos. Quantifying the global parameter tensions between ACT, SPT and Planck. Phys. Rev. D, 103(6):063529, 2021.

[11] Shahab Joudaki et al. KiDS-450: Testing extensions to the standard cosmological model. Mon. Not. Roy. Astron. Soc., 471(2):1259–1279, 2017.

[12] Armin Lederer, Alejandro J Ordóñez Conejo, Korbinian A Maier, Wenxin Xiao, Jonas Umlauft, and Sandra Hirche. Gaussian process-based real-time learning for safety critical applications. In Marina Meila and Tong Zhang, editors, Proceedings of the 38th International Conference on Machine Learning, volume 139 of Proceedings of Machine Learning Research, pages 6055–6064. PMLR, 18–24 Jul 2021.

[13] Armin Lederer, Alejandro Jose Ordonez Conejo, et al. Real-Time Regression with Dividing Local Gaussian Processes. page arXiv:2006.09446, June 2020.

[14] Haitao Liu, Jianfei Cai, Yi Wang, and Yew Soon Ong. Generalized robust Bayesian committee machine for large-scale Gaussian process regression. In Proceedings of the 35th International Conference on Machine Learning, Proceedings of Machine Learning Research, pages 3131–3140. PMLR, 2018.

[15] Janina J. Renk et al. CosmoBit: A GAMBIT module for computing cosmological observables and likelihoods. JCAP, 02:022, 2021.

[16] L. Verde, T. Treu, and A. G. Riess. Tensions between the Early and the Late Universe. Nature Astron., 3:891, 7 2019.

[17] Raymond K. W. Wong, Curtis Byron Storlie, and Thomas C. M. Lee. A frequentist approach to computer model calibration. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 79(2), 5 2016.