technische universität
dortmund

Dissertation zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften

Development of tools for Bayesian data analysis and their application in the search for physics beyond the Standard Model

Cornelius Grunwald

geboren in Bochum

2021

Lehrstuhl für Experimentelle Physik IV
Fakultät Physik
Technische Universität Dortmund

Diese Dissertation wurde der Fakultät Physik der Technischen Universität Dortmund zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften vorgelegt.

Erstgutachter:	Prof. Dr. Kevin Kröninger
Zweitgutachter:	Prof. Dr. Allen Caldwell
Drittgutachter:	Prof. Dr. Gerhard Buchalla
Vorsitz der Prüfungskommission:	Prof. Dr. Joachim Stolze
Vertretung der wiss. Mitarbeiter:	Dr. Bärbel Siegmann

Datum des Einreichens der Arbeit:	30. August 2021
Datum der mündlichen Prüfung:	08. November 2021

Diese Doktorarbeit wurde durch ein Promotionsstipendium der *Studienstiftung des deutschen Volkes* gefördert.

Abstract

In this thesis, methods for Bayesian data analysis are employed in the search for *physics beyond the Standard Model* (BSM). A new implementation of the *Bayesian Analysis Toolkit in Julia* (BAT.jl) is introduced as a modern data analysis framework providing algorithms for Bayesian inference. The EFTfitter.jl package for interpreting measurements in the context of *effective field theories* (EFTs) is presented. It facilitates combining measurements and estimating underlying parameters with Bayesian inference. Both tools are employed for indirect BSM searches using EFTs. Wilson coefficients of dimension-six operators from the top-quark sector of the *Standard Model effective field theory* (SMEFT) are constrained, and approaches for enhancing fits of SMEFT coefficients are investigated. Studies on the effects of correlations between the uncertainties of measurements on the results of SMEFT fits are conducted. It is demonstrated that the correlations can significantly impact the resulting constraints and can be the crucial components deciding whether deviations from the Standard Model of particle physics are observed or not, in particular when assuming future measurements with reduced uncertainties. Moreover, studies on combining measurements of top-quark and flavor physics observables for SMEFT interpretations are performed, and the steps necessary for a combined fit are discussed. Powerful synergies between top-quark processes and $b \to s$ transitions are observed when constraining SMEFT Wilson coefficients of the top-quark sector. It is demonstrated that due to complementary sensitivities, combining observables from different energy scales can tighten the constraints significantly. Future scenarios assuming measurements from HL-LHC, Belle II, and CLIC are investigated, and their potential for improving constraints on up to eleven SMEFT Wilson coefficients is pointed out. The benefits of orthogonal constraints in multidimensional phase spaces for resolving ambiguous solutions are highlighted. The SMEFT studies presented in this thesis demonstrate the capabilities of the new BAT.jl and EFTfitter.jl packages for BSM analyses.

Kurzfassung

In dieser Arbeit werden Methoden der Bayes'schen Datenanalyse bei der Suche nach Physik jenseits des Standardmodells (BSM) eingesetzt. Eine neue Implementation des *Bayesian Analysis Toolkit in Julia* (BAT.jl) wird als modernes Datenanalyse-Framework vorgestellt, welches für Bayes'sche Inferenz benötigte Algorithmen bereitstellt. Ferner wird das Paket EFTfitter.jl für Interpretationen von Messungen im Rahmen von effektiven Feldtheorien (EFTs) vorgestellt. Es ermöglicht die Kombination von Messungen und die Schätzung von Parametern mit Bayes'scher Inferenz. Beide Programme werden für indirekte BSM-Suchen unter Verwendung eines EFT-Ansatzes eingesetzt. Wilson-Koeffizienten von Dimension-Sechs-Operatoren aus dem Top-Quark-Sektor der *Standard Model effective field theory* (SMEFT) werden eingeschränkt und Möglichkeiten zur Verbesserung von Fits der SMEFT-Koeffizienten werden untersucht. Es werden Studien zu den Auswirkungen von Korrelationen zwischen den Unsicherheiten der Messungen auf die Ergebnisse von SMEFT-Fits durchgeführt. Dabei wird gezeigt, dass die Korrelationen einen signifikanten Einfluss auf die resultierenden Einschränkungen haben können und die wesentlichen Komponenten sein können, die darüber entscheiden, ob Abweichungen vom Standardmodell der Teilchenphysik beobachtet werden oder nicht, insbesondere wenn zukünftige Messungen mit reduzierten Unsicherheiten angenommen werden. Darüber hinaus werden Studien zur Kombination von Top-Quark- und Flavor-Physik-Observablen für SMEFT-Interpretationen durchgeführt, und es werden die notwendigen Schritte für kombinierte Fits diskutiert. Starke Synergieeffekte zwischen Top-Quark-Prozessen und $b \to s$ Übergängen für die Einschränkung von SMEFT Wilson-Koeffizienten des Top-Quark Sektors werden beobachtet. Es wird gezeigt, dass aufgrund von sich ergänzenden Sensitivitäten, die Kombination solcher Observablen von verschiedenen Energieskalen die resultierenden Einschränkungen bedeutend verbessern kann. Zukunftsszenarien, die Messungen von HL-LHC, Belle II und CLIC annehmen, werden untersucht, und ihr Potenzial zur Verbesserung der Einschränkungen von bis zu elf SMEFT-Wilson-Koeffizienten wird aufgezeigt. Die Vorteile orthogonaler Einschränkungen in mehrdimensionalen Phasenräumen für das Auflösen von mehrdeutigen Lösungen werden herausgestellt. Die in dieser Arbeit vorgestellten SMEFT-Studien demonstrieren die Leistungsfähigkeit der neu implementierten Pakete BAT.jl und EFTfitter.jl für BSM-Analysen.

Contents

List of contributions

In this thesis, two new tools for Bayesian data analysis are presented and applied in the search for *physics beyond the Standard Model* (BSM) with *effective field theories* (EFTs). Both the development of the tools and the EFT studies have been performed in close collaboration with further scientists and students. Many aspects presented here also have been published as research articles to which I contributed as part of this thesis. In the following, I detail my contributions to the tools and studies presented in this thesis.

- **Chapter 3:** The development of BAT.jl was performed by a team of developers, in particular the authors of Ref. [1]. I am one of the main developers and contributed to several components of the tool. My main contributions were the implementation of the Hamiltonian Monte Carlo algorithm and of the features for output and visualization of results, specifically the plot recipes and the algorithms for estimating highest density regions of marginal distributions. Furthermore, I was deeply involved in implementing the algorithms for low-dimensional sampling and the ellipsoidal nested sampling algorithm. In both cases, I performed pilot studies, provided code, and supervised the bachelor and master students working on these features.

- **Chapter 5:** A first version of the EFTfitter tool was developed by the authors of Ref. [2], to which I contributed as a master's student. A new version of this tool in the Julia programming language, called EFTfitter.jl, is presented here and was developed by myself as part of this thesis. The studies on estimating the weights of individual measurements in their combination with EFTfitter.jl were performed by myself and have not been published before.

- **Chapter 6:** The studies on the impact of correlations on EFT fits presented in this chapter have been published in Ref. [3] and were conducted in collaboration with the other authors of this publication. I was one of the two main analyzers and worked specifically on setting up the EFTfitter.jl analyses and performing the fits and correlation analyses. All plots and results shown in this thesis were derived by myself and are based on the setups developed in Ref. [3]. Additional studies presented in Chap. 6, e.g., on the weights of the measurements and uncertainties, have not been published before.

- **Chapter 7 and Chapter 8:** The studies on the combination of top-quark and B physics observables were conducted in collaboration with the authors of Ref. [4] and Ref. [5]. I was one of the two main analyzers and worked on most aspects of the studies as part of this thesis. In particular, I focused on setting up the EFTfitter.jl analyses, performing the EFT fits, testing the sampling algorithms, and generating the visualizations. I also performed the MC simulations and interpolations to determine the dependence of the fiducial $t\bar{t}\gamma$ cross sections on the Wilson coefficients. All plots and results shown in these chapters have been derived by myself and are based on the setups of Refs. [4, 5].

In the scope of this thesis, I contributed to the following publications as one of the main authors:

- S. Bißmann, J. Erdmann, C. Grunwald, G. Hiller, K. Kröninger, *Constraining top-quark couplings combining top-quark and B decay observables*, Eur. Phys. J. C **80** (2020) 2

- S. Bißmann, J. Erdmann, C. Grunwald, G. Hiller, K. Kröninger, *Correlating uncertainties in global analyses within SMEFT matters*, Phys. Rev. D **102** (2020) 115019

- O. Schulz, F. Beaujean, A. Caldwell, C. Grunwald, V. Hafych, K. Kröninger, S. La Cagnina, L. Röhrig, L. Shtembari , *BAT.jl – A Julia-based tool for Bayesian inference*, SN COMPUT. SCI. **2** 210 (2021) 210

- S. Bißmann, C. Grunwald, G. Hiller, K. Kröninger, *Top and Beauty synergies in SMEFT-fits at present and future colliders*, JHEP **06** (2021) 010

I also presented parts of these works at the *24th International Conference on Computing in High-Energy and Nuclear Physics* (CHEP) 2019 and at the *40th International Conference on High Energy Physics* (ICHEP) 2020 and prepared the following two proceedings:

- A. Caldwell, C. Grunwald, V. Hafych, K. Kröninger, S. La Cagnina, O. Schulz, L. Shtembari, *BAT.jl – Upgrading the Bayesian Analysis Toolkit*, EPJ Web Conf. **245** (2020) 06001

- C. Grunwald, S. Bißmann, J. Erdmann, G. Hiller and K. Kröninger, *Enhancing fits of SMEFT Wilson coefficients in the top-quark sector*, PoS **ICHEP2020** (2021) 323

1 Introduction

The fundamental goal in all fields of research is to generate new knowledge. In modern science, the typical approach for this relies on a combination of two essential components. On the one hand, theoretical descriptions of processes are developed based on the current state of knowledge and novel hypotheses. In physics, such theoretical descriptions are formulated in terms of mathematical models. On the other hand, empirical observations of the processes under study are made by conducting experiments. The crucial step for generating new knowledge is then to compare the model expectations of the current theoretical description with the observed data. In quantitative research, statistical inference is employed to update the state of knowledge in the light of new data. This allows formulating probabilistic statements about the model, its parameters, and its compatibility with the observed data. A formal mathematical approach for this process is provided by Bayesian inference, in which Bayes' theorem is applied to update prior beliefs by taking new data into account. Bayesian inference yields probability distributions of the model parameters, which are straightforward in their interpretation. However, the application of Bayesian methods to real-world problems usually requires the use of sophisticated numerical algorithms, such as *Markov chain Monte Carlo* (MCMC). In this thesis, a modern tool facilitating Bayesian data analysis is presented. The *Bayesian Analysis Toolkit in Julia* (BAT.jl) [1] provides the required infrastructure and a collection of algorithms simplifying the application of Bayesian inference on user-defined problems.

A field of research in which one of the most precise theoretical descriptions of nature has been achieved is particle physics. The *Standard Model of particle physics* (SM) is the underlying theory in this field and has been probed by many experiments. During the development of the SM in the second half of the 20th century, the above-described approach to gaining knowledge through an interplay of theory expectations and experimental data was particularly successful. Theory predictions motivated experimental searches for specific new particles, and, vice versa, unexpected experimental observations led to modifications of the theory. With the discovery of the Higgs boson in 2012 [6, 7], all particles described in the SM have been observed experimentally. Despite this considerable success, there are several aspects, both regarding experimental observations and theoretical details, which indicate that the SM alone cannot be the full theory describing all elementary particles and their interactions. Therefore, particle physics is a highly active field of research, with the current objective being the search for *physics beyond the Standard Model* (BSM). While particle accelerators reached energies above the TeV scale during recent years and a huge amount of data has been analyzed, so far, no direct detection of any BSM phenomena has been achieved. This strengthens assumptions that BSM particles could exist at energy scales much above the scale of the SM and motivates indirect approaches for BSM searches, such as using *effective field theories* (EFTs). EFTs can be applied for model-independent searches for the effects of high-energy BSM phenomena on the lower-energy processes testable at current experiments. The *Standard Model effective field theory* (SMEFT) is a specific EFT extension of the SM and introduces higher-dimensional effective operators into the theory. In the top-quark sector, SMEFT in-

terpretations of measurements gained popularity in recent years [8–23]. They are performed to constrain the coupling strength, so-called *Wilson coefficients*, of the effective operators. In particular, combinations of different observables and various measurements are pursued, aiming towards a global fit of SMEFT coefficients [3–5, 15–23]. The tool EFTfitter.jl [2] builds upon BAT.jl and facilitates such EFT interpretations of measurements. It allows combining possibly correlated measurements and constraining underlying parameters with Bayesian inference. In this thesis, EFTfitter.jl is applied in BSM searches using the SMEFT framework. Emphasis is placed on investigating aspects for enhancing SMEFT fits of multiple observables and measurements constraining Wilson coefficients affecting the top quark. As in such combinations, possible correlations between the uncertainties of measurements are often neglected, the impact of correlations on the results of SMEFT fits is investigated. Moreover, synergies of top-quark and flavor physics observables for constraining SMEFT coefficients of the top-quark sector are explored by performing combined fits. Particular focus is placed on estimating the effects of potential future scenarios with new measurements and reduced uncertainties.

This thesis is organized as follows:
In Chap. 2, an introduction to the main concepts of Bayesian data analysis is given. The difficulties in applying Bayesian inference to multidimensional problems are indicated, and the need for efficient algorithms that allow using Bayes' theorem for statistical inference is motivated. In Chap. 3, BAT.jl is introduced as a tool facilitating Bayesian data analysis. Its design concepts and the included features and algorithms are presented. A brief introduction to the SM and a motivation for searching for BSM physics is given in Chap. 4. The concept of EFTs is introduced, and its application in model-independent BSM searches is motivated. The SMEFT framework and the relevant operators of the top-quark sector of SMEFT are established. In Chap. 5, the combination of measurements for SMEFT fits is motivated, and the concepts for combining possibly correlated measurements with EFTfitter.jl are introduced. Its design structure and features are presented. Studies on estimating the weights of individual measurements in their combination are performed using a toy example. In Chap. 6, the impact of correlations between the uncertainties of measurements on the results of SMEFT fits is investigated systematically. Measurements of single top-quark production and top-quark decay processes are employed to constrain three SMEFT coefficients. Different correlation and uncertainty scenarios are examined and compared. In Chap. 7, a first study on combining top-quark observables with $b \to s$ transitions from B physics in a joint SMEFT fit is performed. The steps necessary for such a combination are introduced, and the Wilson coefficients of three operators affecting the top-quark couplings are constrained. This approach is extended in Chap. 8, where further observables, measurements, and Wilson coefficients are considered in a combined fit to top-quark and flavor physics data. Synergies between the observables from different energy scales are investigated, and their sensitivities on up to eleven SMEFT coefficients are exploited. In particular, possible future scenarios with measurements from HL-LHC, Belle II, and CLIC are examined for their potential to improve the constraints on SMEFT Wilson coefficients. In Chap. 9, the thesis is concluded.

2 Introduction to Bayesian data analysis

The analysis of data is a central task in almost all areas of scientific research and business. Bayesian data analysis provides a systematic approach to gaining knowledge from data. Its main concepts are introduced in this chapter. In Sec. 2.1, certain fundamental concepts of probability theory are described, and Bayes' theorem is derived. Bayesian inference as a rule for updating prior knowledge based on new data is introduced in Sec. 2.2, and specific aspects of the Bayesian interpretation of probabilities are discussed in Sec. 2.3. The difficulties of using Bayesian inference in real-world data analyses are pointed out in Sec. 2.4, and the necessity of efficient numerical algorithms is motivated. The concept of *Markov chain Monte Carlo* (MCMC) methods as a solution to these problems is introduced in Sec. 2.5, and the Metropolis–Hastings algorithm as the original and best-known MCMC algorithm is presented.

2.1 Probabilities and Bayes' theorem

Early mathematical approaches to the treatment of probabilities emerged mainly in the 17th and 18th centuries [24]. The first axiomatic foundations for the concept of probability were formulated in 1933 by Kolmogorov in terms of set theory [25]. Since these three Kolmogorov axioms are the most widely accepted conceptual definition and do not imply any specific interpretation of probability, they are briefly stated here. [1]

Axiom 1 – Non-negativity The probability p of each possible event E in a sample space S is a real number greater than or equal to zero,

$$p(E) \geq 0 \quad \forall \quad E \in S\,. \tag{2.1}$$

Axiom 2 – Unit measure The probability p that at least one of all possible events in the sample space S will occur is one,

$$p(S) = 1\,. \tag{2.2}$$

Axiom 3 – Additivity If two events A and B are mutually exclusive, i.e., $A \cap B = \emptyset$, the probability p of either A or B occurring is the sum of the individual probabilities of A and B,

$$p(A \cup B) = p(A) + p(B)\,. \tag{2.3}$$

[1] Kolmogorov's axioms represent the first but not the only possible axiomatic definition of the concept of probability. In particular for the Bayesian interpretation of probabilities, the theorems formulated by Cox [26] in terms of conditional probabilities are an often discussed alternative. Kolmogorov's axioms can be derived from them [27, 28].

Following these fundamental postulates, all further rules for the calculus of probabilities can be derived, including, for example, the rule of monotonicity, i.e., $p(A) \leq p(B)$ if $A \subseteq B$, or the complement rule, $p(A^C) = 1 - p(A)$, where A^C is the complement of A.

An important concept is that of conditional probabilities, which allows formulating the probability $p(A|B)$ for A to occur under the prerequisite that B has already occurred. In the formalism laid out by Kolmogorov's axioms, conditional probabilities are described as the quotient of the probability of both events to happen, i.e., the intersection $A \cap B$, divided by the probability of B to happen,

$$p(A|B) = \frac{p(A \cap B)}{p(B)} \,. \tag{2.4}$$

Bayes' theorem When multiplying Eq. (2.4) by the denominator $p(B)$ of the right side, an expression for the joint probability of A and B both occurring is obtained,

$$p(A \cap B) = p(A|B)\, p(B) \,. \tag{2.5}$$

Using the equality $p(A \cap B) = p(B \cap A)$ and again dividing by $p(B)$, the famous equation

$$p(A|B) = \frac{p(B|A)p(A)}{p(B)} = \frac{p(B|A)p(A)}{\sum_i p(B|A_i)p(A_i)} \tag{2.6}$$

is obtained, which is commonly referred to as *Bayes' theorem*. The last equality in Eq. (2.6) applies when the sample space S consists of multiple disjoint partitions S_i. Bayes' theorem is named after Thomas Bayes (c. 1701-1761), who was the first to write about a mathematical approach to inferring from an effect back to its cause, which he did by solving a specific inverse problem through a learning process. Bayes' manuscripts were edited and published posthumously by Richard Price in 1763 [29]. It was, however, Pierre-Simon Laplace, who (during the late 18th and early 19th century) developed a thorough mathematical treatment for conditional probabilities and brought the concepts of Bayes and Price into the form of Eq. (2.6) that is nowadays known as Bayes' theorem [30].

2.2 Bayesian inference

The central idea of Bayesian inference is to use Bayes' theorem for updating knowledge about a set of parameters in the light of new data. For this purpose, Bayes' theorem in Eq. (2.6) is rewritten for model-based inference as

$$p(\vec{\theta}|\vec{D}, M) = \frac{p(\vec{D}|\vec{\theta}, M)\, p_0(\vec{\theta}|M)}{P(\vec{D}|M)} \,. \tag{2.7}$$

In this formulation, $\vec{\theta}$ are the parameters of a statistical model M, and $\vec{D}$ are the observed data. The term $p(\vec{D}|\vec{\theta}, M)$ describes the probability of observing the data $\vec{D}$, given the parameters $\vec{\theta}$ and the model M. As in Bayesian inference the parameters $\vec{\theta}$ are varied for fixed data $\vec{D}$,

this term is not necessarily a valid probability density in $\vec{\theta}$ and is, therefore, referred to as the *likelihood*. The term $p_0(\vec{\theta}|M)$ is the probability distribution of the model parameters $\vec{\theta}$ before the data $\vec{D}$ are taken into account. It is called the *prior probability distribution*, or *prior* for short. The desired result of the inference process is $p(\vec{\theta}|\vec{D}, M)$, the so-called *posterior probability density*, often simply called the *posterior*. It is the probability distribution of the model parameters $\vec{\theta}$ when taking into account both the observed data $\vec{D}$ and the prior knowledge about the parameters $\vec{\theta}$. The denominator of Eq. (2.7),

$$p(\vec{D}|M) = Z = \int p(\vec{D}|\vec{\theta}, M)\, p_0(\vec{\theta}|M)\ \mathrm{d}\vec{\theta}\,, \tag{2.8}$$

is called *marginal likelihood* or *evidence*.[2] It is a normalizing constant which ensures that the posterior is a valid probability density. It can be interpreted as the marginal probability of observing the data $\vec{D}$ when assuming the model M. It is discussed in Sec. 2.4 that, in particular, this denominator complicates the application of Bayesian inference and that an explicit computation of the evidence is usually avoided. In Bayesian model comparison, however, the evidence plays a central role. The so-called *Bayes factor* is given as

$$K = \frac{p(\vec{D}|M_A)}{p(\vec{D}|M_B)} = \frac{Z_A}{Z_B}\,, \tag{2.9}$$

which is the ratio of the marginal likelihoods of two different models, M_A and M_B, describing the same data $\vec{D}$. The value of the Bayes factor K can be used as a criterion for model selection, with values $K > 1$ supporting M_A over M_B and values $K < 1$ rather supporting M_B.

Bayes' theorem in Eq. (2.7) describes a knowledge update process in which prior beliefs about the distribution of model parameters, $p_0(\vec{\theta}|M)$, are updated according to observed data, $p(\vec{D}|\vec{\theta}, M)$, resulting in an updated belief on the distribution of the parameters, $p(\vec{\theta}|\vec{D}, M)$. This process is the central concept of data analysis with Bayesian inference. Particularly in cases where access to new data is gained gradually, Bayesian inference allows a stepwise update of the knowledge since it is possible to use the posterior of a previous analysis as the prior of the next analysis with additional data. The resulting posterior is then, however, exactly the same as when analyzing all data in a single step.

When performing Bayesian inference, it is often the posterior distribution of an individual parameter that is of interest. It can be obtained through marginalization, i.e., by integrating over all parameters that are not of interest,

$$p(\theta_i|\vec{D}, M) = \int p(\vec{\theta}|\vec{D}, M) \prod_{j \neq i} \mathrm{d}\theta_j\,. \tag{2.10}$$

In typical applications, the statistical model M depends not only on the parameters of interest $\vec{\theta}$, but also on a number of nuisance parameters $\vec{\nu}$. In such cases, Bayes' theorem takes the

[2] The formulation of the evidence as an integral is valid for continuous probability distributions. Bayes' theorem also applies for discrete distributions, for which the integral in the denominator is replaced by a sum.

form

$$p(\vec{\theta}, \vec{\nu}|\vec{D}, M) = \frac{p(\vec{D}|\vec{\theta}, \vec{\nu}, M)\, p_0(\vec{\theta}, \vec{\nu}|M)}{P(\vec{D}|M)}\,. \tag{2.11}$$

Information about the parameters of interest is obtained by marginalization over the nuisance parameters,

$$p(\vec{\theta}|\vec{D}, M) = \int p(\vec{\theta}, \vec{\nu}|\vec{D}, M)\ \mathrm{d}\vec{\nu}\,. \tag{2.12}$$

2.3 Bayesian interpretation of probabilities

While Kolmogorov's axioms provide a foundation for the mathematical treatment of probabilities, they do not imply any interpretation of the concept of probability itself. Over the years, different interpretations have been developed and widely disputed. While some approaches are claimed to be *objective*, others are considered more *subjective* interpretations. The most famous example of the former class is the so-called *frequentist* interpretation, in which a probability is thought of as the relative frequency of a certain outcome in the limit of infinite repetitions of the same process. The Bayesian interpretation, on the other hand, is typically considered a subjective approach since probability is interpreted as a *degree-of-belief*. Because different statistical methods for inference are associated with the respective interpretations of probability, there are discussions about the choice of methods and the validity of corresponding results. Especially in the case of the Bayesian and frequentist interpretations, there are often contentious and almost philosophical disputes about the validity or superiority of one interpretation over the other, which will not be discussed here but have been addressed, for example, in Refs. [31–36]. However, it should be noted that not only the statistical methods themselves but also their results and the interpretability thereof depend on the interpretation of probability chosen and can even be contradictory, as, for example, demonstrated in the Jeffreys–Lindley paradox [37, 38].

Bayes theorem itself, as given in Eq. (2.6), is a universal formula for probabilities that is valid for all interpretations, not only for Bayesian reasoning. When using this theorem for statistical inference, as shown in Eq. (2.7), however, the Bayesian interpretation is adopted. This is particularly noticeable in the required formulation of prior knowledge (or prior beliefs) in terms of prior distributions. The presence of a prior is one of the main aspects that distinguish the Bayesian approach from other statistical methods. While a prior distribution allows to include existing knowledge about the parameters, it is often the researchers' subjective choice how exactly this information is formulated. This subjectivity leads to the prior often being the target for criticism of Bayesian methods. As the prior can indeed have a significant influence on the outcome of the inference, it is best practice to explicitly state the prior distributions chosen. Comparison studies using different prior distributions can be performed to investigate the influence of the prior on the actual analysis. There are also approaches aiming towards more *objective Bayesian analyses*, for example, by minimizing the information contained in the prior using concepts of information theory. Different methods have been developed to construct

so-called *uninformative* prior distributions, with the *Jeffreys prior* and the *reference priors* being the two most famous objective Bayesian approaches [39–42].

One of the main advantages of Bayesian inference, compared to other approaches, is that the outcome of the analysis is the full probability distribution of the model parameters. Based on the posterior probability, it is possible to derive summary statistics and make probabilistic statements that are usually straightforward in their interpretation. For example, the credible intervals derived from the posterior distribution of a Bayesian analysis are generally more informative and easier to interpret than the confidence intervals determined in frequentist analyses. However, obtaining a distribution as the result of the inference is also one of the aspects that complicates the application of Bayesian inference, in particular, when using multivariate models with many free parameters, as is discussed in the next section.

2.4 Difficulties in the application of Bayesian inference

Even though the concept of Bayesian inference in Eq. (2.7) seems straightforward, its application for data analysis usually requires numerical computations with efficient algorithms. This is because analytical calculations of the posterior distribution are only possible in a very limited number of cases. In particular, the evidence integral in Eq. (2.8) is usually intractable, especially when considering models with multiple parameters. Thus, the posterior distribution typically cannot be determined in a closed-form but has to be approximated or represented by samples. Also, it is usually not only the posterior density itself that is of interest, but rather certain summary statistics of the distribution, such as the mean value, the variance, or other moments and interval estimates. The computation of such estimates then again requires integrations that are analytically intractable.

In particular, for inference problems with many parameters, instead of trying to solve the integrals, methods for sampling from the unnormalized posterior distribution are employed. The desired information about the properties of the posterior distribution is obtained from samples, using the *law of large numbers*. Therefore, sampling or approximating the posterior distribution is a central task when applying Bayesian inference. However, efficiently sampling an arbitrary target distribution becomes increasingly difficult as the number of dimensions increases since the volume of the parameter space grows exponentially with the number of free parameters. For example, when sampling a one-dimensional parameter space, simple quadrature-based algorithms divide the continuous sample space into a number of bins, N. In d dimensions, however, this approach requires dividing the parameter space into N^d hypercubes. Obviously, this approach quickly becomes infeasible as the number of required hypercubes grows exponentially with the number of dimensions. This effect is often referred to as the *curse of dimensionality* and makes the application of numerical methods in high-dimensional parameter spaces challenging.

Furthermore, when sampling distributions with many parameters, it has to be considered that in high dimensions, most of the probability mass is no longer distributed around the mode, even though this is where the probability density is highest. Since probability is the integral

over the probability density, also the corresponding volumes have to be taken into account. In the simple picture of the parameter space subdivided into hypercubes, the probability is the product of the density and the corresponding cube volumes. In high dimensions, however, most of the volume is not centered around the mode but rather in the tails of the distribution. Therefore, for efficient sampling it must be ensured that samples are drawn from the region where the probability mass is high, which is called the *typical set* [43, 44]. The typical set has the important property that samples drawn from this region represent the target distribution faithfully, which makes the sampling efficient. In the example of the N^d hypercubes, this means that the evaluations of most hypercubes are unnecessary since they are not part of the typical set and contribute almost nothing to the total probability. While for low-dimensional problems, typically for $d \leq 5$, it is still feasible to perform a simple grid-based posterior evaluation, more efficient approaches for sampling of posterior distributions are needed for models with a larger number of parameters.

Monte Carlo (MC) methods, which use random numbers to explore the posterior distribution, are a common approach to sampling. However, simple rejection-based MC algorithms are inefficient for sampling from the typical set. Approaches like *importance sampling*, where samples are generated not from the unknown posterior distribution but from a similar distribution that can be sampled more efficiently, require some prior knowledge about the shape of the posterior distribution and typically also scale poorly with the number of parameters.

2.5 Markov chain Monte Carlo

The technique that revolutionized the application of Bayesian inference for data analysis is MCMC. The concept of MCMC was developed in 1953 and was originally designed for physics simulations [45]. It was not until the 1990s that MCMC gained popularity in statistical analyses and was first used for Bayesian inference [46]. By providing a method for efficiently sampling posterior distributions, even in cases with many parameters, MCMC made the application of Bayesian inference to real-world problems feasible. Together with the steep increase in available computing power in recent decades, the development of MCMC methods led to a major boost in the relevance of Bayesian analyses. Today, MCMC algorithms are the most common approach for performing Bayesian inference, and their main concepts are briefly introduced here.

A *Markov chain* is a sequence of states $\{X_i\}$ that satisfies the *Markov property*, which is the requirement that the transition probability $p(X_{i+1}|X_i)$ from the current state X_i into the next state X_{i+1} only depends on the current and next state but not on any of the previous states. If the conditional probability distribution for the transition $p(X_{i+1}|X_i)$ does not depend on i,

$$p(X_{i+1}|X_i) = \pi \;\forall\; i\,, \tag{2.13}$$

the distribution π is called the *stationary* or *equilibrium* distribution of the Markov chain. If such a stationary distribution exists, starting from any initial point in the state space, this distribution will eventually be reached. When using Markov chains for MC sampling, the goal

is to construct a Markov chain in such a way that its stationary distribution is the distribution of interest, i.e., the *target distribution*. In Bayesian inference, this target distribution is the posterior distribution. In practice, the most challenging tasks are constructing a Markov chain that converges quickly to its equilibrium distribution and determining whether or not the chain has converged to this distribution. When a chain has reached convergence, the states X_i visited by the chain are samples of the target distribution. The collection of states generated by the Markov chain before converging to its stationary distribution is commonly referred to as the *warm-up* or *burnin* phase. Samples generated in this phase are usually discarded since they do not represent the target distribution accurately. Determining the convergence of Markov chains is one of the most important tasks when using MCMC methods, as only the samples of a converged chain should be used to derive information about the target distribution. Common criteria for determining the convergence of Markov chains are introduced in Sec. 3.3.2.

A drawback of MCMC methods that has to be considered is that the samples generated by Markov chains are not completely independent but are correlated. Autocorrelations arise since each state depends on the previous one. When constructing Markov chains, the goal is to reduce such autocorrelations in order to approximate the target distribution as efficiently as possible. Often, the autocorrelation is taken into account by referring to the so-called *effective sample size* (ESS), which describes the number of *independent and identically distributed* (iid) samples that would be equivalent to the number of correlated MCMC samples.

2.5.1 Metropolis–Hastings algorithm

The first algorithm for generating Markov chains with a specific stationary distribution was the Metropolis algorithm [45], which was originally used for sampling the Boltzmann distribution. It was later generalized to be used for arbitrary target distributions by Hastings [47] and brought into the form that is now the most common algorithm for MCMC and is widely used in Bayesian inference.

The Metropolis–Hastings algorithm works as follows:
Starting from an initial point θ^i in the parameter space, a new point $\tilde{\theta}^i$ is proposed according to a proposal function $g(\tilde{\theta}^i|\theta^i)$. This proposed point is then either accepted as the next point of the chain, or it is rejected. If it is rejected, the chain stays at the point θ^i, and a new proposal point is generated. The probability to accept a proposed point in the accept-reject step is given as

$$p_{\text{accept}} = \min\left(1, \frac{\pi(\tilde{\theta}^i)}{\pi(\theta^i)} \frac{g(\theta^i|\tilde{\theta}^i)}{g(\tilde{\theta}^i|\theta^i)}\right) , \tag{2.14}$$

where π is the target distribution that needs to be evaluable at any point of the parameter space. In Algorithm 1, the Metropolis–Hastings algorithm is illustrated in terms of pseudocode.

Algorithm 1 Metropolis–Hastings

start at a random initial point in the parameter space: θ^i
for i in 1:N **do**
    propose a new point: $\tilde{\theta}^i = \texttt{rand}(g(\theta|\theta^i))$
    calculate the acceptance ratio $w = \frac{\pi(\tilde{\theta}^i)}{\pi(\theta^i)} \cdot \frac{g(\theta^i|\tilde{\theta}^i)}{g(\tilde{\theta}^i|\theta^i)}$
    draw a random number from a uniform distribution: $u = \texttt{rand}(0, 1)$
    if $w \leq u$ **then**
        accept the proposal and move to the proposed point: $\theta^{i+1} = \tilde{\theta}^i$
    else
        reject proposal and stay at current point: $\theta^{i+1} = \theta^i$
    end if
end for
return all points $[\theta^1, \ldots, \theta^N]$

A special property of the Metropolis–Hastings algorithm is that it generates Markov chains that fulfill the *detailed balance* principle. A Markov chain that satisfies detailed balance is reversible, i.e., the probability of the chain moving from state X_i to state X_j is the same as the probability of moving from X_j to X_i,

$$\pi_i \, p(X_j|X_i) = \pi_j \, p(X_i|X_j) \, . \tag{2.15}$$

The detailed balance principle is a stronger condition than that required for the existence of a stationary distribution. Thus, it is ensured that Markov chains generated by the Metropolis–Hastings algorithms have a stationary distribution that is the target distribution.

While the basic ideas of the Metropolis–Hastings algorithm are straightforward, achieving implementations that are efficient for arbitrary target distributions can be challenging. To reach a fast convergence of the Markov chains, the proposal functions usually have to be tuned, and adaptive steps have to be introduced into the sampling process. A tool providing such a general-purpose implementation of the Metropolis–Hastings and further sampling algorithms is introduced in the next chapter.

3 BAT.jl – A modern toolkit for Bayesian data analysis

The concept of Bayesian inference introduced in the previous chapter is a powerful approach to data analysis. But as already indicated, its application to real-world problems requires the use of efficient numerical algorithms for sampling and integration. In this chapter, the *Bayesian Analysis Toolkit* (BAT), a software tool providing algorithms facilitating the use of Bayesian inference for user-defined problems, is introduced. The main ideas of BAT are motivated in Sec. 3.1. In Sec. 3.2, a new version of the toolkit in the Julia programming language, called BAT.jl, is presented, and important design considerations and aspects of its implementation are discussed. The numerical algorithms included in BAT.jl are described in Sec. 3.3, and features for the output and visualization of analysis results are presented in Sec. 3.4.

3.1 The Bayesian Analysis Toolkit

As motivated in Sec. 2.4, performing data analysis with Bayesian methods usually requires advanced numerical techniques, such as MCMC. While the general concept of MCMC algorithms, such as the Metropolis–Hastings algorithm presented in Sec. 2.5.1, is straightforward, achieving performant and reliable implementations of such methods that are universally applicable can be challenging. This is, for example, due to the presence of tunable hyperparameters and many computational details. Therefore, it is clearly expedient to use tools that provide elaborate and well-tested implementations of the required algorithms. A variety of such tools exists for the realm of Bayesian analyses [48–51]. In the field of particle and high-energy physics, BAT [52] is an established framework for Bayesian analyses. Integrating into the common infrastructure of this field of research, BAT was initially implemented in the C++ programming language and relying on the particle-physics analysis framework ROOT [53]. Since its first release in 2008, the main focus of BAT was to provide a user-friendly infrastructure that allows performing Bayesian analyses with user-specified likelihoods and prior distributions. This is achieved by offering a collection of algorithms for posterior exploration, optimization, and integration, facilitating typical analyses tasks such as parameter estimation, model comparisons, and hypothesis testing. Particular focus was placed on providing an efficient implementation of the Metropolis–Hastings algorithm, including an adaptive tuning procedure for the proposal distribution [52].

Over the years, BAT has been used in various analyses in particle physics, for example, for model fitting [54–56] or limit setting [57–60]. BAT also had use cases in related fields of research, such as cosmology [61], astrophysics [62], and nuclear physics [63]. In order to open the tool for further areas of application, its dependencies on domain-specific software had to be reduced, and the computational requirements of modern data analysis frameworks had to be met. For these reasons, a completely new implementation of BAT was targeted.

3.2 BAT.jl – A modern implementation of BAT

Starting in 2017, BAT was rewritten from scratch in the Julia programming language. The first stable version of the new software package, called *Bayesian Analysis Toolkit in Julia* (BAT.jl), was released in 2019 and has been significantly updated in its second version in 2020 [1]. Julia is a modern multi-purpose programming language with particular focus on high-performance numerical computing [64]. The choice of the Julia language for the development of BAT.jl allows for performant implementations of the numerical algorithms required for Bayesian analyses while providing a high-level and user-friendly code syntax. Because of its inherent support of parallel and distributed computing, as well as its features allowing automatic code-based differentiation, Julia is a well-suited basis that allows meeting the requirements of a modern toolkit for Bayesian data analysis.

BAT.jl is a registered Julia package, which makes its installation and setup very simple. Installation and update processes can be performed automatically using Julia's integrated package management system.

```
Pkg.add(BAT)
using BAT
```

The design of BAT.jl is focused on (normalized or unnormalized) probability densities. This allows formulating problems in terms of user-specified likelihood and prior distributions without requiring the use of a certain modeling or domain-specific language. The posterior density is constructed from the custom likelihood and prior definitions.

```
posterior = PosteriorDensity(likelihood, prior)
```

Similar to its predecessor, BAT.jl aims to provide a large collection of algorithms for common tasks in Bayesian analyses, focusing on capable general-purpose implementations of sampling, integration, and optimization algorithms. Intuitive high-level functions facilitate the application of the numerical algorithms needed for Bayesian inference.

```
samples  = bat_sample(posterior, sampler).result
mode     = bat_findmode(posterior, estimator).result
evidence = bat_integrate(samples, integrator).result
```

The numerical algorithms available in BAT.jl are introduced in Sec. 3.3. For all algorithms, reasonable default settings are implemented, which depend on the properties of the considered objects and can be adjusted according to the users' needs. For the prior, constrained or unconstrained distributions can be used, typically provided by the `Distributions.jl` package [65]. The use of hierarchical priors is also supported. Since many algorithms perform best in certain parameter spaces, or even require a specific parameter space, such as the unit-hypercube, BAT.jl includes features for automatic parameter-space transformations.

```
trafo = PriorToUniform()
```

BAT.jl offers a variety of algorithms for sampling, integration, and optimization so that the user can choose the algorithms that best suit the problem at hand. In the following, the numerical algorithms included in BAT.jl and some important details of their implementations are introduced.

3.3 Numerical algorithms of BAT.jl

3.3.1 Algorithms for low-dimensional sampling

As discussed in Sec. 2.4, for low-dimensional problems (typically $d \leq 5$), it can be computationally feasible to evaluate the posterior distribution at a certain number of discrete points in the phase space that are not adjusted to the target density. By assigning weights corresponding to the value of the posterior distribution at these points, independent samples can be obtained. Even though this approach suffers from the curse of dimensionality and does not scale well with an increasing number of dimensions, three algorithms of this type are included in BAT.jl for low-dimensional use cases. As such algorithms do not rely on tunable hyperparameters and are independent of the specific posterior shape, the results obtained with these methods do not suffer from biases or convergence problems and always represent the true target distribution. Only the resulting resolution of the posterior distribution depends on the number of sampling points chosen. Therefore, these algorithms provide important cross-checks when comparing with the results from more sophisticated MCMC sampling algorithms.

The simplest strategy for generating sampling points in the d-dimensional phase space is using an evenly spaced grid on a constrained parameter space. BAT.jl includes an algorithm for performing such a grid-based posterior evaluation.

```
sampler = GridSampler(ppa = 100)
```

The number of evaluation points per parameter can be specified, and the total number of samples generated is ppa^d. In addition, BAT.jl includes a similar algorithm that is based on the Sobol sequence [66].

```
sampler = SobolSampler(nsamples = 10^6)
```

Using the Sobol sequence, quasi-random numbers are obtained that are evenly distributed in the d-dimensional hyperspace and have low discrepancy. This algorithm allows generating an arbitrary number of samples. Compared to the `GridSampler` it has the advantage that the sampling points are not equidistantly spaced. The last sampling algorithm of this type in BAT.jl generates sampling points by drawing random numbers from the prior. This is possible because prior distributions in BAT.jl are required to be directly samplable, yielding iid samples.

```
sampler = PriorImportanceSampler(nsamples = 10^6)
```

3.3.2 MCMC algorithms

Burn-in and convergence

As mentioned in Sec. 2.5, one of the most important steps when using MCMC algorithms for Bayesian inference is to ensure that the Markov chains have converged towards their target distribution so that the samples appropriately represent the posterior distribution. When using MCMC algorithms in BAT.jl, a pre-run is performed during which hyperparameters of the sampling algorithms are tuned, and the convergence of the chains is checked iteratively.

```
burnin = MCMCMultiCycleBurnin()
```

For the pre-run, multiple Markov chains are initialized with random starting points drawn from the prior. The pre-run is performed in cycles during which the chains are iterated for a number of steps (a certain percentage of the requested samples by default) and tuned according to algorithm-specific tuning procedures. The proposal function is tuned to optimize the acceptance rate of proposed sampling points and to improve the convergence speed. After each tuning cycle, the convergence of the chains is checked using the convergence criteria introduced in the next paragraph. The pre-run is stopped when all chains have converged and are tuned to the required acceptance rate. The samples generated during the pre-run are discarded, and the desired number of posterior samples is generated by further iterating the converged chains. The resulting samples of the posterior distribution are returned. If not all of the chains converge within the specified number of tuning cycles, a warning is displayed, and the user can decide in advance whether the sampling should be continued or aborted in this case.

Convergence criteria

Two common criteria for checking the convergence of Markov chains are implemented in BAT.jl. They are both based on a comparison of the variance of samples within a chain (within-chain variance W) to the variance of samples between different chains (between-chain variance B), as proposed by Gelman and Rubin [67]. If both variances are compatible, it indicates that the chains have mixed well and that the samples of each chain have converged to the same stationary distribution. Following Ref. [67], this is mathematically checked by calculating the so-called *potential scale reduction factor* (PSRF) as

$$\hat{R} = \sqrt{\frac{W + \frac{1}{n}(B - W)}{W}}\,, \tag{3.1}$$

where n is the number of samples per chain. In BAT.jl, a cutoff value of $\hat{R} \leq 1.1$ is used by default to determine convergence. The Gelman–Rubin approach has been generalized for the multivariate case by Brooks and Gelman [68], which is the default convergence criterion in BAT.jl.

```
convergence = BrooksGelmanConvergence(threshold=1.1)
```

Metropolis–Hastings algorithm

The default sampling algorithm for posterior distributions in BAT.jl is the Metropolis–Hastings algorithm, as introduced in Sec. 2.5.1.

```
sampler = MCMCSampling(mcalg=MetropolisHastings(), nchains=4, nsteps=10^6)
```

A multivariate Students' t-distribution is used as the default proposal function. As mentioned above, during the pre-run multiple Markov chains are generated. In the tuning process of the Metropolis–Hastings algorithm, the acceptance rate α, i.e., the ratio of accepted proposals to the total number of samples in each chain, is adapted to be within a range $\alpha_{\text{min}} \leq \alpha \leq \alpha_{\text{max}}$ to ensure an optimal exploration and performance of the chain. By default, $\alpha_{\text{min}} = 0.15$ and $\alpha_{\text{max}} = 0.35$ are chosen. The acceptance rate is tuned using an adaptive procedure that updates the covariance matrix of the proposal distribution based on the covariance matrix of the samples generated in the previous tuning cycle. This matrix is then scaled with an overall factor to reach the desired acceptance rate.

Hamiltonian Monte Carlo

Hamiltonian Monte Carlo (HMC), formerly also referred to as *Hybrid Monte Carlo*, is another prominent MCMC-based sampling algorithm [43, 69, 70]. It aims to circumvent the often inefficient random walk behavior of the Metropolis–Hastings algorithm by utilizing gradient information of the posterior to adjust the proposals to the shape of the target distribution. As a consequence, HMC usually yields higher acceptance rates than the Metropolis–Hastings algorithm and reaches convergence faster. The samples obtained with HMC tend to be less correlated and, therefore, yield a higher ESS, reducing the total number of samples required for exploring the target distribution appropriately. HMC samplers are capable of working in up to several thousand dimensions. However, these advantages come with the cost that a single step of the Markov chain is computationally more expensive in HMC than in the Metropolis–Hastings algorithm. A brief introduction to the main concept of HMC is given here. A more detailed overview of HMC techniques can be found in Refs. [43, 70], for example.

The central idea of HMC is to expand the d-dimensional parameter space to $2d$ dimensions by introducing hyperparameters $\vec{p}$, the so-called momenta. The original phase space is thus transformed into the canonical phase space $\vec{q} \rightarrow (\vec{q}, \vec{p})$.[1] Also the target distribution $\pi(\vec{q})$ is lifted into the canonical phase space by constructing a joint distribution

$$\pi(\vec{q}, \vec{p}) = \pi(\vec{p}|\vec{q})\pi(\vec{q}) = \mathrm{e}^{-H(\vec{q},\vec{p})}\,, \tag{3.2}$$

where a conditional probability distribution $\pi(\vec{p}|\vec{q})$ is chosen for the momenta.

[1] In this paragraph, the model parameters to be estimated with Bayesian inference are exceptionally denoted as $\vec{q}$ instead of $\vec{\theta}$ in order to follow the common notation for HMC.

In the last equality of Eq. (3.2), the so-called *Hamiltonian*

$$H(\vec{q},\vec{p}) = -\log\pi(\vec{q},\vec{p}) = -\log\pi(\vec{p}|\vec{q}) - \log\pi(\vec{q}) \tag{3.3}$$

is introduced. With this definition of the Hamiltonian, it is possible to use the methods of classical mechanics, where *Hamilton's equations of motion* are given as

$$\frac{\mathrm{d}q_i}{\mathrm{d}t} = \frac{\partial H}{\partial p_i}, \qquad \frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial H}{\partial q_i}\,. \tag{3.4}$$

When using HMC, a new point $(\vec{q}^{\,*},\vec{p}^{\,*})$ in the canonical phase space is obtained by solving the equations of motion for a certain time T. This is achieved numerically using so-called *symplectic integrators* [71, 72] and allows to move along trajectories ϕ in the canonical phase space, resulting in a transition

$$(\vec{q},\vec{p}) \to \phi_T(\vec{q},\vec{p}) = (\vec{q}^{\,*},\vec{p}^{\,*})\,. \tag{3.5}$$

A new proposal point $\vec{q}^{\,*}$ in the original parameter space is then obtained by marginalizing over the momenta $\vec{p}$. Similar to the Metropolis–Hastings algorithm, the new proposal is then either accepted or rejected by calculating an acceptance ratio.

While this approach for generating proposals seems straightforward, achieving efficient and reliable implementations of HMC algorithms can be delicate. Compared to the Metropolis–Hastings algorithm, HMC requires several additional steps, such as the transformations between the canonical phase space and the original parameter space, or the numerical integration of the equations of motions. This also leads to HMC sampling processes being dependent on multiple tunable hyperparameters, such as the momenta $\vec{p}$, the step sizes, and number of steps used in the symplectic integrators. In particular, HMC also requires information about the gradient of the posterior distribution, which has to be provided by the user or needs to be automatically derivable via code-based differentiation. Different approaches have been developed to find optimal values for the tuning parameters of HMC [73]. Adaptive techniques like the *No-U-Turn sampler* (NUTS) [74] have proven to be successful by eliminating the need for certain tuning parameters and are, therefore, the current state-of-the-art approaches for HMC.

In BAT.jl, HMC is provided using an internal interface to the `AdvancedHMC.jl` package [75] for performing the individual HMC sampling steps.

```
sampler = MCMCSampling(mcalg=AdvancedHMC(), nchains=4, nsteps=10^6)
```

`AdvancedHMC.jl` provides robust and efficient implementations of HMC and allows choosing different techniques for the symplectic integrations and generation of proposals. Specifically, multiple versions of NUTS are available and can be selected. The high-level operations, such as the construction and transformation of the posterior distribution, as well as the burn-in process, are handled by BAT.jl, similar to the Metropolis–Hastings sampling. The required gradient of the target distribution can often be derived automatically from the user-defined

posterior distribution by making use of Julia's efficient support of automatic differentiation, for example, via the packages `ForwardDiff.jl` [76] and `Zygote.jl` [77]. In cases where this is not possible, the gradient can also be specified explicitly by the user.

3.3.3 Integration algorithms

The sampling algorithms discussed above do not take into account the integral in the denominator of Bayes' theorem in Eq. (2.7) and sample only the unnormalized posterior distribution. To estimate the evidence Z, however, for example, when aiming to calculate Bayes' factors for model comparisons, it is necessary to evaluate the corresponding integral. In BAT.jl, different approaches for the integration of posterior distributions are included.

An interface to the CUBA library [78] allows estimating integrals with established general-purpose integration algorithms based on importance sampling techniques, such as VEGAS [79].

```
evidence = bat_integrate(posterior, VEGASIntegration()).result
```

As these algorithms do not scale very well with the dimensionality, a new algorithm for integration based on samples generated with the sampling algorithms described above, is included in BAT.jl, called *Adaptive Harmonic Mean Integration* (AHMI) [80].

```
evidence = bat_integrate(samples, AHMIntegration()).result
```

The AHMI algorithm creates multiple hyperrectangles in the parameter space and computes the harmonic mean estimator based on the posterior samples in each of these regions. By joining the results from all hyperrectangles, an integral estimate for the full posterior distribution is obtained. Furthermore, estimates for the uncertainty are provided. This algorithm has been shown to perform well for up to 20 dimensions [80].

3.3.4 Partitioned sampling

Using the integration algorithms introduced above, a new sampling approach is implemented in BAT.jl, which allows a parallel exploration of the sampling space. The main idea of this so-called *partitioned sampling* [81] is to divide the parameter space into multiple subspaces and to perform the posterior sampling separately in each of them, using the sampling algorithms discussed above. When merging the samples from the individual subspaces, the relative weights of the samples have to be taken into account. For this purpose, the samples of each subspace are first reweighted according to their integral, determined using the AHMI algorithm. This allows for correctly stitching together all subspaces.

```
sampler = PartitionedSampling(npartitions=4, sampler=mcmc, integrator=ahmi)
```

The partitioned sampling allows reducing wall-clock times by providing another option for parallelizing the sampling process. Also, for certain target distribution, such as multimodal

posterior distributions, the partitioned sampling approach is advantageous, since the included integration step makes the correct estimation of the relative heights of the individual modes possible, which can not be achieved using the ordinary Metropolis–Hastings algorithms, as will be demonstrated later in Sec. 8.5.

3.3.5 Nested sampling

A MC approach that yields both samples of the target distribution and an estimate of its integral, i.e., the evidence Z in Bayesian inference, is nested sampling. The idea of nested sampling was proposed in 2004 [82], and its main concepts for the exploration of the phase space and the estimation of the integral are briefly summarized below, following the review of Ref. [83].

When using nested sampling for Bayesian inference, a number N of so-called *live points* is randomly drawn from the prior distribution. Each of these points is associated with a fraction $1/N$ of the total volume, and the likelihood is evaluated at each of them. The live point with the lowest likelihood value, L_1, is then removed, which results in a shrinkage of the volume by $\delta V = 1/N$ and leads to a remaining volume of $V = (1 - \frac{1}{N})$. The removed point (*dead point*) is saved as a sample and weighted by the product of its likelihood value and the corresponding volume. A new live point is then sampled from the prior distribution, with the requirement of having a likelihood value greater than that of the removed point, i.e., greater than L_1. This is referred to as *constrained sampling* or *likelihood-restricted prior sampling* (LRPS). With the new point, again N live points are available, and each of them represents a fraction $1/N$ of the remaining volume. The steps of removing the live point with the lowest likelihood value and sampling a new live point from the prior distribution with increasing likelihood values are repeated iteratively. This causes the volume to shrink by a constant factor in each iteration, and the considered phase space becomes more and more concentrated around the highest likelihood values. The nested sampling algorithm terminates when the changes in the remaining volume after i iterations, $V_i = (1 - \frac{1}{N})^i$, become negligibly small. The evidence is then estimated as

$$Z \approx \sum_i \Delta V_i \cdot L_i \,, \tag{3.6}$$

where $\Delta V_i = V_i - V_{i-1} = \frac{1}{N}(1 - \frac{1}{N})^i$ is the shrinkage of volume, and L_i are the likelihood values of the corresponding dead points. The removed live points are samples of the target distribution. The crucial step in nested sampling is generating new live points that satisfy the requirement of increasing likelihood values, i.e., the LRPS. Various approaches have been proposed for this constrained sampling, also including methods based on HMC approaches [84], for example. A detailed overview and comparison of such algorithms for nested sampling is given in Ref. [83].

In BAT.jl, two different implementations of nested sampling algorithms are available using interfaces to specialized packages. On the one hand, an interface to the python package

`UltraNest`[2] is provided by the Julia wrapper `UltraNest.jl`[3].

```
sampler = ReactiveNestedSampling(min_num_live_points=400)
```

The `UltraNest` package implements the *MLFriends* algorithm [85, 86] for the constrained sampling, which uses ellipsoids around the live points to draw new points from these regions. It has the advantageous property of being free of tuning parameters. On the other hand, a native Julia implementation of nested sampling is accessible using an internal interface to the `NestedSamplers.jl` package[4], which provides single- and multi-ellipsoidal nested sampling algorithms [87].

```
sampler = EllipsoidalNestedSampling(num_live_points=400)
```

Besides the ability to compute posterior samples and the integral estimate simultaneously, nested sampling algorithms are particularly suited for exploring multimodal distributions, as will be demonstrated in Chap. 8.

3.3.6 Optimization and best fit values

A quantity of particular interest in Bayesian analyses is the point with the highest value of the posterior probability, i.e., the global mode of the posterior distribution. When samples of the posterior distribution are already available, the sample with the highest weight can be a first estimate for the global mode. For more precise estimates of the global mode, BAT.jl includes two optimization algorithms by interfacing to the `Optim.jl` package [88]. The default optimization algorithm for mode finding in BAT.jl is the Nelder-Mead algorithm [89].

```
estimator = MaxDensityNelderMead()
```

For distributions for which the gradient is known or can be derived automatically, the LBFGS algorithm [90] can be applied.

```
estimator = MaxDensityLBFGS()
```

Often, also the mode values of marginal distributions, further referred to as *marginal modes*, are of interest. In BAT.jl, the marginal mode is estimated from binned samples as the center of the bin with the highest weight. The number of bins to be used for the calculations can be set manually or determined automatically using different binning rules, such as the Freedman-Diaconis rule [91], for example.

[2] `https://johannesbuchner.github.io/UltraNest/`
[3] `https://github.com/bat/UltraNest.jl`
[4] `https://github.com/TuringLang/NestedSamplers.jl`

3.4 Visualization of results

An essential aspect of each data analysis is the presentation and visualization of the results. BAT.jl provides several functionalities for easy access to numerical and graphical output of results. A summary object can be constructed containing information about the sampled distribution, such as estimates for the mean, global and marginal modes, and the covariance matrix.

```
results = SampledDensity(posterior, samples)
```

An example of the output obtained from this object is shown in Fig. A.1 of App. A. All properties and estimates can also be accessed via specific functions to use them for further computations or to create custom output formats of the results. For the one-dimensional marginal distributions, additional summary statistics, such as the quantiles or smallest intervals containing a certain probability, can be determined.

For the graphical representation of results, BAT.jl includes multiple plot recipes that facilitate plotting with the `Plots.jl` package [5]. For the visualization of high-dimensional distributions, by default, the one- and two-dimensional marginals are plotted using histogrammed samples.

```
plot(samples)
```

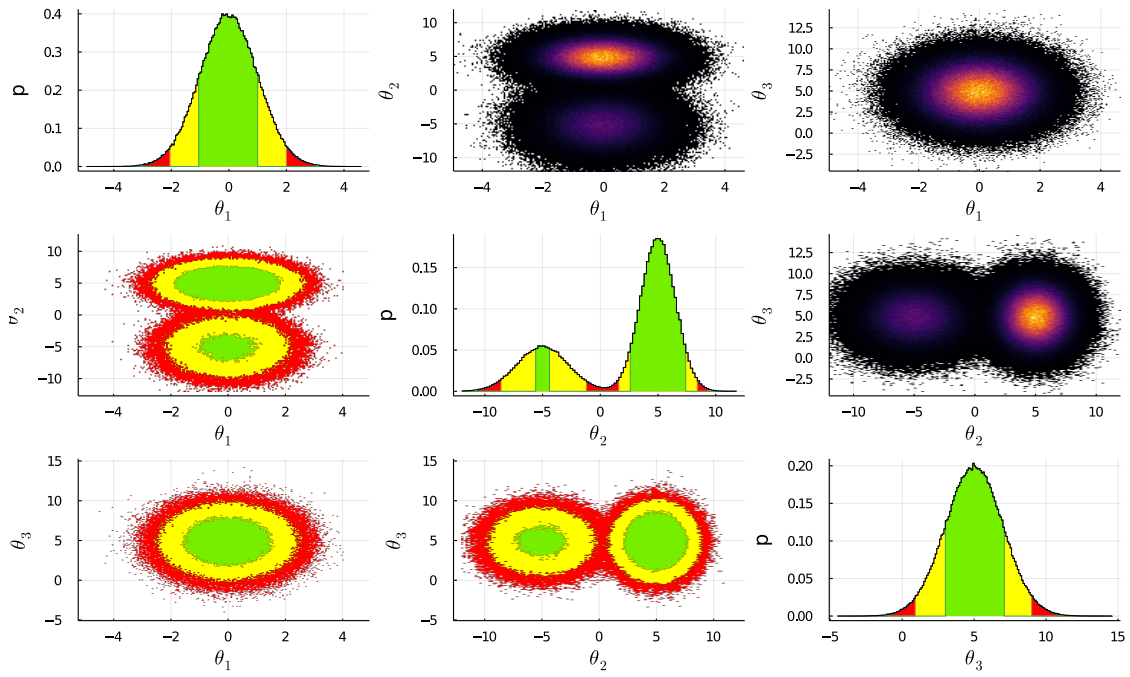

Figure 3.1: Example of a BAT.jl default plot, showing one- and two-dimensional marginal distributions. For the one-dimensional marginals on the diagonal and the two-dimensional marginals in the lower triangle, the smallest intervals (areas) containing 68.3 %, 95.5 % and 99.7 % of the posterior probability are highlighted in green, yellow, and red, respectively. In the upper triangle plots, two-dimensional heatmaps are shown.

[5] `https://github.com/JuliaPlots/Plots.jl`

In Fig. 3.1, a BAT.jl default plot is depicted for a three-dimensional multimodal distribution, built from multiple normal distributions. By default, the smallest intervals (areas in the two-dimensional case) containing 68.3 %, 95.5 % and 99.7 % of the posterior probability are highlighted. For the one-dimensional distributions, the central intervals, i.e., the symmetrical intervals around the mean, can also be highlighted instead. The corresponding intervals and areas are determined based on histogrammed samples. A more detailed discussion of this approach is given in Sec. 5.4.

Various keyword arguments allow for highlighting certain properties of the distribution and for customizing the plotting styles. By plotting the posterior distribution together with the associated prior distribution, the knowledge update gained in the Bayesian inference can be visualized, as demonstrated in Fig. 3.2.

```
p = plot(samples, :θ2, nbins=300, mean=true, intervals=[0.5, 0.7, 0.9],
    colors=blues, marginalmode=Dict("linecolor"=>:red, "linewidth"=>2))
p = plot!(prior, :θ2, linecolor=:darkorange2, linewidth=2)
```

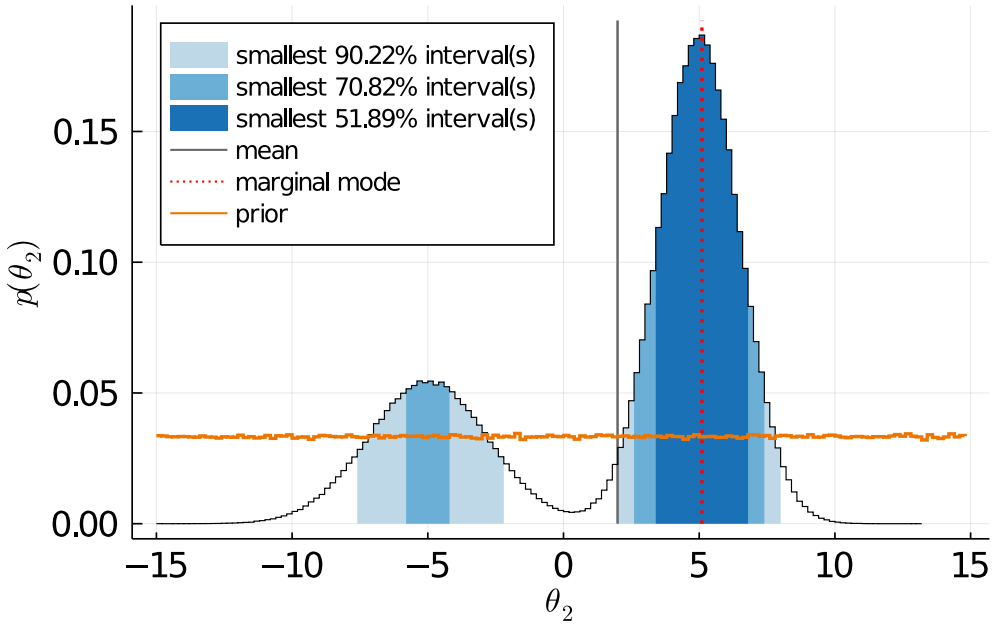

Figure 3.2: Example of a customized plot showing both the marginal posterior and the prior distribution. The marginal mode and the mean of the posterior distribution are shown. The smallest intervals containing 50 %, 70 % and 90 % of the posterior probability are highlighted.

Since full access to the posterior samples and the marginalized histograms is available, completely customized visualizations of the results can also be created.

4 The search for physics beyond the Standard Model

In particle physics, the statistical analysis of large amounts of data plays a crucial role. In this thesis, the previously presented methods and tools for Bayesian data analysis are applied in the search for *physics beyond the Standard Model* (BSM) using the framework of *effective field theories* (EFTs). This approach is motivated in this chapter. For this purpose, a short introduction to the *Standard Model of particle physics* (SM) is given in Sec. 4.1. Open questions and aspects that motivate the search for BSM phenomena are presented in Sec. 4.2. The concept of EFTs is introduced in Sec. 4.3, and the *Standard Model effective field theory* (SMEFT) as a model-independent approach to BSM searches is discussed.

4.1 A brief introduction to the Standard Model

The grand objective of particle physics is to search for the fundamental components of the universe and to understand the interactions between them. The SM comprises the current state of knowledge about the elementary particles and their interactions and is the underlying model of particle physics. The theory of the SM and the corresponding mathematical description were developed gradually during the 20th century and finalized in its current form in the 1970s [92–100]. Within the SM, all elementary particles discovered so far and three of the four known fundamental forces are described, with only the gravitational force not being included in its formalism. For the purposes of particle physics, however, the gravitational force is orders of magnitudes weaker than the other three forces, and its influence can be neglected. The mathematical formulation of the SM allows making predictions for the values of observable physical quantities, which can then be probed by particle physics experiments.

The elementary particles described in the SM are grouped into two main classes according to their spin values. The so-called *fermions* have half-integer spin, while *bosons* have integer-valued spin.

Fermions The Fermions are the constituents of matter. They are divided into two subclasses, called *quarks* and *leptons*. The six *quarks* carry fractions of the elementary electric charge e. They are grouped into the up-type quarks named up (u), charm (c), and top (t) quark with an electric charge of $+\frac{2}{3}e$, and the down-type quarks called down (d), strange (s), and bottom/beauty (b) quark with an electric charge of $-\frac{1}{3}e$. The leptons are also grouped according to their electric charge. The electron e^-, the muon μ^-, and the tau lepton τ^- have an electric charge of $-1e$ and are referred to as the charged leptons. For each of the charged leptons, there exists a corresponding uncharged lepton, called *neutrino*. Specifically, these are the electron neutrino (ν_e), the muon neutrino (ν_μ), and the tau neutrino (ν_τ), which are assumed to be massless in the SM. For both the quarks and the leptons, a mass hierarchy

between the three respective generations is observed, with the particles of the third generation being the heaviest and those of the first generation being the lightest. The top quark is the heaviest particle of the SM, with a mass of about $(172.76 \pm 0.30)\,\mathrm{GeV}$ [101]. For each of the fermions, there is a corresponding antifermion with opposite charges, forming the *antimatter*. Only for the neutrinos, it is not clear if there exist separate antineutrino particles or if the neutrinos themselves are their own antiparticles, which would make them so-called *Majorana* particles [101].

Bosons The bosons are grouped into the *gauge bosons* with spin one and the scalar *Higgs boson* with spin zero. The gauge bosons are the mediators of forces. By coupling to specific charges, they enable interactions between correspondingly charged particles. The photon γ is the gauge boson of the electromagnetic interaction. It couples to the electric charge and therefore to all fermions, except for the neutrinos, and to the charged bosons. The gluon g is the mediator of the strong interaction and couples to the *color charge*, a quantum number carried by the quarks and also the gluons themselves. This special property of the gluons being color-charged allows for self-interactions between them. In ordinary phases of matter, quarks and gluons cannot be isolated because of the so-called *confinement*, which describes that only color-neutral combinations of these particles can be observed. Together with the fact that the coupling strength of the strong force increases with increasing distances between the particles, this leads to the *hadronization*, in which new particles are generated from the vacuum to form color-neutral bound states of quarks and gluons, the *hadrons*. The gauge bosons of the weak interaction are the charged $W^{\pm}$ and the neutral Z^0 bosons. In contrast to the photon and the gluons, the weak gauge bosons are massive, with masses of $m_W = (80.379 \pm 0.012)\,\mathrm{GeV}$ and $m_Z = (91.187 \pm 0.002)\,\mathrm{GeV}$ [101]. The weak interaction is related to the quantum number called *weak isospin* and is the only interaction that affects all fermions. In 2012, the Higgs boson H was experimentally discovered as the last fundamental particle described in the SM [6, 7]. It arises from the Higgs mechanism [98–100], which describes how the charged fermions and massive bosons obtain their mass through a spontaneous breaking of the electroweak symmetry.

Mathematical formulation of the SM Mathematically, the SM is formulated as a relativistic and renormalizable quantum field theory with a $\mathrm{SU}(3)_C \otimes \mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ gauge symmetry group before *spontaneous symmetry breaking* (SSB). The $\mathrm{SU}(3)_C$ is the underlying symmetry of *quantum chromodynamics* (QCD), i.e., the theory describing the strong interaction. A unified description of *quantum electrodynamics* (QED), i.e., the theory of the electromagnetic interaction, and the weak theory is achieved in the *electroweak theory*, which has the $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ gauge symmetry [92, 93, 96]. The Higgs mechanism induces a SSB of the electroweak symmetry and gives rise to the three massive bosons. After this symmetry breaking, the gauge group of the SM is $\mathrm{SU}(3)_C \otimes \mathrm{U}(1)_{\mathrm{EM}}$.

A mathematical summary of the SM quantum field theory can be given in terms of a Lagrangian density, often simply called *Lagrangian*. From the Lagrangian, certain properties of the particles and their interactions, as well as rules for calculating predictions of observable

processes, can be derived. The Lagrangian is built from quantum fields and corresponding operators. All terms of the SM Lagrangian are of mass-dimension four, except for the $\varphi^\dagger\varphi$ part in the Higgs mass term, which is of dimension two. A compact version of the SM Lagrangian is given in Eq. (4.1), together with a brief description of the respective terms:

$$\begin{aligned}\mathcal{L}_{\text{SM}} = &\underbrace{-\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}G^A_{\mu\nu}G^{A\mu\nu} - \frac{1}{4}W^I_{\mu\nu}W^{I\mu\nu}}_{\text{kinetic terms of gauge bosons}} + \underbrace{(D_\mu\varphi)^\dagger(D^\mu\varphi)}_{\text{kinetic Higgs term}} + \underbrace{\mu^2\varphi^\dagger\varphi - \lambda(\varphi^\dagger\varphi)^2}_{\text{Higgs mass term}} \\ &+ \underbrace{i(\bar{u}_R \not{D} u_R + \bar{d}_R \not{D} d_R)}_{\text{kinetic quark terms}} + \underbrace{i(\bar{l}_L \not{D} l_L + \bar{e}_R \not{D} e_R + \bar{q}_L \not{D} q_L)}_{\text{kinetic lepton terms}} \\ &- \underbrace{(Y^u \bar{q}_L \tilde{\varphi} u_R + Y^d \bar{q}_L \varphi d_R + Y^e \bar{l}_L \varphi e_R + \text{h.c.})}_{\text{mass terms for quarks and charged leptons}} . \end{aligned} \tag{4.1}$$

In this notation, $B_{\mu\nu}$, $W_{\mu\nu}$, and $G^A_{\mu\nu}$ are the gauge field strength tensors of the $\mathrm{U}(1)_Y$, $\mathrm{SU}(2)_L$, and $\mathrm{SU}(3)_C$ gauge groups. The Higgs doublet is denoted as φ, where $\tilde{\varphi} = i\sigma_2\varphi^*$ with the Pauli matrices σ_i. The $\mathrm{SU}(2)_L$ doublets for quarks and leptons are $q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}$ and $l_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}$. The terms u_R, d_R, and e_R are the singlets for the up-type quarks, down-type quarks, and charged leptons, respectively. The Y^i denote the Yukawa matrices, and D_μ is the covariant derivative.

4.2 Motivation for physics beyond the Standard Model

The SM is a huge success of particle physics, as it describes all known particles and their interactions comprehensively and allows making predictions for measurable quantities. Predictions based on the SM have been tested successfully by various particle physics experiments over a large range of energy scales [101]. Until now, no deviations between measurements and corresponding SM predictions have been found that would be considered significant at the $5\,\sigma$ level, which is the common significance threshold in particle physics. There are, however, several indications that the SM is not the full theory describing all fundamental components and interactions of the universe. On the one hand, there are experimental observations of phenomena that cannot be explained within the SM or that deviate in a consistent way from the SM predictions. On the other hand, there are certain shortcomings of the theoretical framework of the SM. A non-exhaustive selection of such indications is listed below.

Neutrino masses Multiple experiments observed so-called *neutrino oscillations*, which are conversions of the neutrino flavors from one lepton generation into another. Such conversions have been observed using atmospheric [102], solar [103], reactor [104], and accelerator [105] neutrinos. The mixing between different neutrino generations is theoretically described by the Pontecorvo–Maki–Nakagawa–Sakata matrix [106, 107], which requires the neutrinos to be massive. In the SM, however, neutrinos are considered to be massless, and it is not possible to add terms for neutrino masses to the Lagrangian while keeping the same particle content and

symmetries of the SM. There are approaches that allow adding mass terms for the neutrinos, such as the *see-saw mechanism* or the concept of *sterile neutrinos*, which are two examples of possible extensions of the SM [101].

Dark matter & dark energy Cosmological observations of the rotation curves of galaxies [108] or of gravitational lensing effects [109] imply the presence of matter in a form that differs from the ordinary particles described by the SM and that is not directly detectable with present particle physics experiments. Current estimates based on the Planck satellite data [110] suggest that only 5 % of the universe is made of the baryonic matter described in the SM. Around 27 % are estimated to be matter of a still unknown nature, so-called *dark matter*. The remaining 68 % percent are an unknown form of energy, called *dark energy*. That the baryonic matter described by the SM makes up only a small fraction of all matter indicates that the SM alone cannot be the full theory describing all particles of the universe.

Baryon asymmetry The aforementioned 5 % of observable matter in the universe are assumed to consist entirely of baryonic matter. So far, no hints for primordial forms of antibaryonic matter have been found. Antimatter has only been observed in particle accelerators or cosmic rays, in both cases resulting from the collisions of matter particles [101]. Within the SM, however, there is no mechanism preferring baryonic matter over antibaryonic matter. The absence of observable antimatter in the universe is referred to as the *baryon asymmetry*, for which the SM does not provide an explanation.

***B* anomalies** Recent measurements of different observables in semi-leptonic B meson decays deviate from their SM predictions, e.g., Refs. [111–116], hinting towards a violation of lepton flavor universality, a fundamental principle of the SM. The deviations observed in the individual measurements are at a level of up to $3.6\,\sigma$ [116] and are therefore not yet considered statistically significant on their own. However, these deviations are observed in multiple observables at different experiments and are all pointing in the same direction. Therefore, the so-called B *anomalies* are currently a highly interesting sector for searching deviations from the SM. Further experimental results will need to be considered to confirm or disprove the observed anomalies and possible violations of lepton flavor universality.

Anomalous magnetic moments Due to higher-order contributions, the SM expectations for the magnetic moments of charged fermions deviate from the value $g = 2$ expected from the Dirac equation. The so-called anomalous magnetic moment $a = (g - 2)/2$ is a quantity describing this deviation, and its value is subject to high-precision experiments. A deviation between the measured value of a and the corresponding SM predictions could be an indication that contributions from unknown particles are neglected in the calculations. For the anomalous magnetic moment of the muon, a_μ, a deviation from the SM prediction was observed at the Brookhaven National Laboratory [117]. Recent results from Fermilab confirm this deviation, with the combined significance of both experiments now being $4.2\,\sigma$ [118]. As this value is

still below the common threshold of $5\,\sigma$, further experimental results will have to be considered. Because the computation of the SM prediction for a_μ is intricate, in particular, due to contributions from hadronic light-by-light scattering, also improved SM predictions have to be considered, which might decrease the currently observed deviations between theory and experiment [119].

Unification of forces The SM provides consistent descriptions for three of the four known fundamental forces in terms of quantum field theories. While the electromagnetic and weak forces can be unified in the electroweak theory, a combined description including also the strong force would be desirable and is often referred to as a *grand unified theory*. A description of gravity as a quantum field theory has not been achieved so far. From a theoretical point of view, however, a common framework describing all known interactions in a coherent way, i.e., a so-called *theory of everything*, would be desirable for particle physics.

The points above, as well as further aspects not listed here, indicate that the SM alone is not the full theory of particle physics, even though it describes a huge part of empirical observations very precisely. Therefore, there is still a lot of research ongoing in the field of particle physics, both in experiment and in theory, searching for observations and theoretical descriptions of BSM physics.

4.3 Going beyond the Standard Model with effective field theories

In the search for BSM physics, various approaches are being pursued. One approach is to develop hypothetical BSM models postulating specific new particles or phenomena, which are then directly searched for by the experiments. On the other hand, there are indirect searches probing only the effects of potential BSM scenarios on known processes by comparing observations with the SM predictions. One such indirect and model-independent approach is the use of EFTs for BSM searches.

4.3.1 The concept of effective theories

Effective theories can be applied when a separation of energy scales exists. In such cases, it is possible to formulate an effective description of processes at the low energy scale of interest without knowing the full fundamental theory defined at a higher energy scale. The effective description then includes only the relevant effects at the scale of interest and ignores effects at other scales. The use of this concept of separated energy (or length) scales is well known in various fields of research. For example, in atomic physics and chemistry, precise calculations of observable processes are possible without taking into account the substructure of protons and neutrons since this becomes only relevant at much higher energy scales, which are not probed in these research fields [120, 121].

In particle physics, effective theories can be applied, for example, if the energy scale of a mediator particle is much higher than the scale of the process under study. The most famous

example of such an effective description in particle physics is Fermi's theory of beta decay. In his publication from 1934, Fermi describes the beta-minus decay, $n \to p + e^- + \bar{\nu}$, as a four-fermion point interaction, which agrees well with observations [122]. In the 1960s, however, the full theory of the weak interaction was developed by Glashow, Salam, and Weinberg [92, 93, 96], and the W boson was postulated to be the particle mediating the weak interaction in the beta decay. It was discovered in 1983 by the UA1 and UA2 experiments [123, 124]. Fermi's theory thus turned out to be only an effective description of the beta decay. However, due to the large mass of the W boson, compared to the energy scale of the beta decay itself, Fermi's theory still provides a good description of the process, even though it is unaware of the W boson exchange as the actual underlying mechanism. Considering the weak interaction and its mediator particles to be the full theory, valid also for higher energies, Fermi's theory for processes at lower energy scales can be recovered. Such a procedure, starting from the known fundamental theory and deriving an effective description for lower scales, is referred to as the *top-down* approach and is demonstrated in the following for the example of the beta decay.

In Fig. 4.1, two Feynman diagrams for the beta-minus decay of a neutron are depicted, neglecting spectator quarks. The left diagram shows the process as described by the weak interaction, mediated by a W boson. The right diagram shows the same process as a four-point interaction, as described by Fermi.

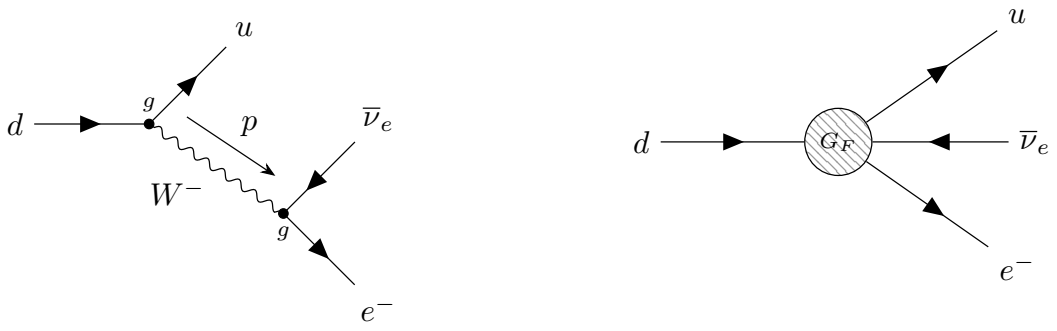

Figure 4.1: Feynman diagrams for the beta-minus decay as described by the weak interaction (left) and by Fermi's effective theory (right).

In beta decays, the energy scale of the process, i.e., the transferred momentum p, is much smaller than the energy scale of the full theory, which corresponds to the mass of the mediator, m_W. Therefore, the propagator of the massive W boson in the left diagram of Fig. 4.1 can be expanded as

$$\frac{g^2}{p^2 - m_W^2} = -\frac{g^2}{m_W^2}\left[1 + \frac{p^2}{m_W^2} + \mathcal{O}\left(\frac{p^4}{m_W^4}\right)\right] \approx -\frac{g^2}{m_W^2}\,. \tag{4.2}$$

Thus, starting from the weak interaction, a low-energy effective description is obtained, and the Fermi coupling constant $G_\mathrm{F} \sim \frac{g^2}{m_W^2}$ can be recovered. Since the heavy W boson is integrated out in this low-energy approximation, the Feynman diagram takes the form of a four-point contact interaction with the Fermi constant G_F as the coupling strength, as is shown in the right diagram of Fig. 4.1.

When applying such an effective approach while being aware of the energy scale of the underlying full theory, the inaccuracy Δ_{EFT} introduced by using the effective description can be estimated by comparing the energy scale E of the process to the energy scale Λ of the full theory. Considering the above example of the beta-minus decay of the neutron, an estimate of the EFT error can be approximated as

$$\Delta_{\text{EFT}} \sim \frac{E^2}{\Lambda^2} \sim \frac{m_n^2}{m_W^2} \approx 10^{-4}\,. \tag{4.3}$$

For a process with similar Feynman diagrams but an even lower energy scale, such as the muon decay $\mu \to e\bar{\nu}_e\nu_\mu$, the corresponding error of the Fermi approach would be only $\Delta_{\text{EFT}} \sim \frac{m_\mu^2}{m_W^2} \approx 10^{-6}$ [121]. It is thus apparent that, if the energy scales are separated well enough, effective theories can yield very precise descriptions even though heavy mediator particles are not taken into account. For energies E close to the scale Λ, however, the EFT expansion breaks down and no longer provides a valid description of the process.

In B physics, EFT approaches similar to Fermi's theory are commonly applied for the computation of low-energy observables, in particular for flavor-changing processes. Since such processes take place at energies well below the electroweak scale, a suitable EFT can be derived by integrating out the particles above this scale, i.e., the massive $W^\pm$, Z, and Higgs bosons, as well as the top quark and hypothetical heavy BSM particles. The resulting effective theory is referred to as the *weak effective theory* (WET) [125–127] and is commonly used for the calculation of observables. A complete basis of effective operators below the electroweak scale has recently also been formulated in the *low-energy effective field theory* (LEFT) [128]. In Chap. 7 of this thesis, the WET is introduced in more detail and employed to compute BSM contributions to B physics observables.

4.3.2 The Standard Model Effective Field Theory

In the search for BSM physics, experiments at current particle accelerators are able to directly search for new phenomena in an energy range of up to a few TeV. However, even though the integrated luminosity of analyzed data has grown rapidly, no signs of BSM phenomena have been detected in such direct searches. A possible reason for this could be that the energy scale of potential BSM physics might be much higher than the scale of the SM, and therefore not reachable with current accelerator experiments. This would, however, imply a separation of the energy scales of SM and BSM physics. The SM itself could then be considered a low-energy effective description of a more general theory at a much higher energy scale, often referred to as an *ultraviolet* (UV) complete theory. In the search for such a full theory of SM and BSM physics at an energy scale Λ much higher than that of the SM, a so-called *bottom-up* EFT approach can be pursued. By systematically extending the SM Lagrangian, it is attempted to parametrize the effects of new phenomena in terms of effective operators, without having specific knowledge about the details of the full theory. By constraining the coupling strengths of effective operators, BSM physics can be probed in a model-independent way by searching for deviations from the SM expectations. However, the full UV theory at the scale Λ cannot be

determined with this approach, as the same effective operators could be derived from different fundamental theories having similar low-energy effects. Conversely, when considering a specific UV theory at scale Λ, however, it can be matched to the effective theory.

A common effective extension of the SM for BSM searches is SMEFT. In SMEFT, it is assumed that the energy scale of BSM physics Λ is well above the vacuum expectation value of the Higgs boson, i.e., $v = 246\,\text{GeV}$, and that the Higgs field is a doublet in the SU(2) symmetry [129].[1] The Lagrangian of SMEFT is built by adding higher-dimensional terms to the SM Lagrangian. For this, d-dimensional operators are constructed from the SM fields and are required to be gauge invariant under the symmetries of the SM. As all terms of the SM Lagrangian are of mass dimension four at most, the SMEFT operators of dimensions $d > 4$ are suppressed by inverse powers of the BSM scale, $1/\Lambda^{(d-4)}$. For each effective operator O, there is a corresponding *Wilson coefficient* C, which quantifies its coupling strength. The Lagrangian of d-dimensional SMEFT operators is therefore given as

$$\mathcal{L}^{(d)} = \sum_i \frac{1}{\Lambda^{d-4}} C_i^{(d)} O_i^{(d)} \,, \tag{4.4}$$

where i iterates over all possible operators of dimension d. The SMEFT Lagrangian is then created by adding all higher-dimensional terms to the SM Lagrangian

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \sum_{d=5}^{\infty} \mathcal{L}^{(d)} \,. \tag{4.5}$$

For $d = 5$, only one effective operator can be constructed in SMEFT. This operator does not conserve lepton number and gives rise to Majorana neutrino masses [130]. In general, only even-dimensional operators conserve lepton and baryon numbers [131, 132]. As the higher-dimensional terms are suppressed by increasing powers of Λ, the leading contributions to BSM physics (besides the neutrino masses) come from dimension-six operators. In this thesis, only these leading BSM contributions, suppressed by Λ^{-2}, are considered, and the SMEFT Lagrangian under study reads

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \sum_i \frac{1}{\Lambda^2} C_i^{(6)} O_i^{(6)} + \mathcal{O}\left(\frac{1}{\Lambda^4}\right) . \tag{4.6}$$

For $d = 6$, it has been shown that there are 59 independent effective operators when counting only one fermion generation [130, 133]. All operators are given in Ref. [133] in a representation referred to as the *Warsaw basis*, which is also employed in this thesis. When considering all possible flavor combinations of the fermion fields, these 59 operators lead to a total of 2499 independent parameters for dimension six. In observable processes, however, usually only a small subset of these operators gives non-negligible contributions [130, 131]. This allows deriving constraints on the Wilson coefficients of the operators from measurements.

[1] There are other EFT approaches to BSM physics, such as the *Higgs effective field theory* (HEFT), in which the Higgs boson is assumed to be a SU(2) singlet [129].

Based on Eq. (4.6), predictions for the values of observables can be calculated. Similar as for the SM, the Lagrangian is translated into Feynman rules, which are then applied to Feynman diagrams. When computing the Feynman amplitude $\mathcal{M}$ of a specific process, all corresponding Feynman diagrams need to be considered, including those containing modifications from EFT operators. In this thesis, only Feynman diagrams with a single insertion of a dimension-six operator are considered, as contributions from diagrams with multiple insertions are suppressed by more than Λ^{-2}. The resulting amplitude can be divided into contributions from SM diagrams and from BSM diagrams including EFT operators [134],

$$\mathcal{M} = \mathcal{M}^{\text{SM}} + \frac{1}{\Lambda^2}\sum_i C_i \mathcal{M}_i^{\text{BSM}} + \mathcal{O}\left(\frac{1}{\Lambda^4}\right) . \tag{4.7}$$

Cross sections are proportional to the square of the absolute value of the amplitude, i.e., $\sigma \propto |\mathcal{M}|^2$. The resulting terms when squaring Eq. (4.7) are given as

$$\sigma = \sigma^{\text{SM}} + \frac{1}{\Lambda^2}\sum_i C_i \sigma_i^{\text{interf.}} + \frac{1}{\Lambda^4}\sum_{i\leq j} C_i C_j \sigma_{ij}^{\text{BSM}} + \mathcal{O}\left(\frac{1}{\Lambda^6}\right) . \tag{4.8}$$

The first term, σ^{SM}, contains only the contributions from the SM diagrams. The leading contributions to BSM physics come from the interference between SM and BSM diagrams and are of the order $\mathcal{O}(\Lambda^{-2})$. Contributions of the order $\mathcal{O}(\Lambda^{-4})$ arise from squaring the amplitude of the BSM diagrams. The origin of these terms is also visualized in Fig. 4.2.

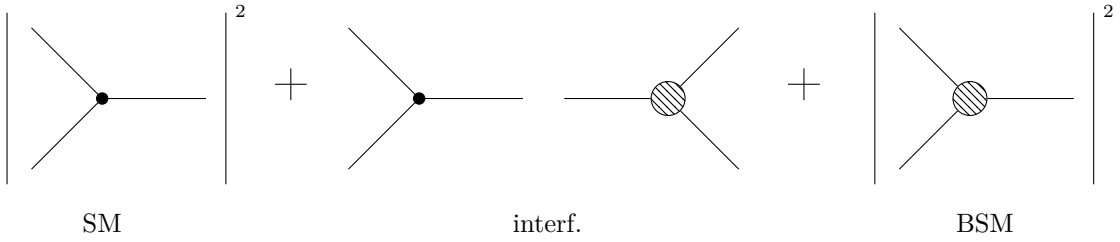

Figure 4.2: Diagrams visualizing the origin of the terms in Eq. (4.8). Adapted from [134].

In this thesis, Eq. (4.8) is used to obtain computationally fast parametrizations of observables as a function of the Wilson coefficients. This is achieved by determining the values of the terms $\sigma_i^{\text{interf.}}$ and σ_{ij}^{BSM} in Eq. (4.8) through interpolations to MC simulations performed with fixed values of the Wilson coefficients. Terms of the order $\mathcal{O}(\Lambda^{-4})$ in Eq. (4.8) are in general not considered to be leading order BSM contributions for dimension-six operators. However, in some processes, the interference terms of certain operators might be suppressed, for example, due to small SM contributions. In such cases, the leading BSM contributions from dimension-six operators can be those of $\mathcal{O}(\Lambda^{-4})$.[2] It has been shown that, for these reasons, the corresponding BSM terms coming from squaring the contributions of diagrams with dimension-six operators should not be neglected in general [134]. In this thesis, observables, such as cross

[2] Contributions of the same order also arise from the interference of dimension-eight effective operators with the SM. Such contributions are not taken into account in this thesis since dimension-eight operators are not considered.

sections and decay widths, are therefore usually parametrized using Eq. (4.8), including the quadratic terms of the order $\mathcal{O}\left(\Lambda^{-4}\right)$. For the EFT studies in Chap. 6, the relevance of the quadratic terms is investigated explicitly by comparing both a linear parametrization taking into account only the interference terms and a quadratic parametrization taking also into account the contributions of order $\mathcal{O}\left(\Lambda^{-4}\right)$ [15].

4.3.3 The top-quark sector of SMEFT

This thesis focuses on SMEFT operators affecting the top quark, as it plays an important role in the search for BSM physics. In many BSM scenarios, new phenomena are expected to be best accessible in top-quark processes. This is motivated by certain unique properties of the top quark. The top quark is the heaviest particle of the SM and has the largest Yukawa coupling to the Higgs boson, being of the order $\mathcal{O}(1)$. This makes studies of the coupling between the top quark and the Higgs boson very interesting and sensitive to potential new phenomena. In addition, many BSM scenarios assume that new particles decay into SM particles, with the top quark often taking a special role in these hypothetical decays. Due to its high mass, also the average lifetime of the top quark is very small, being at the order of $10^{-25}\,\mathrm{s}$ [101, 135]. This time is so short that the top quark decays before hadronization. This allows studying the top quark and its properties as a quasi-free particle, which is a unique feature in the quark sector. As its lifetime is also shorter than the timescale for spin decorrelation, the top quark keeps its spin state from production to decay and passes its spin completely to the decay products [101, 135]. Therefore, studies of top-quark decays, which happen almost exclusively into a W boson and a b quark, are of particular interest. For these reasons, BSM studies using top quarks are very common, and during the recent years EFT interpretations of top-quark measurements gained a lot of interest and importance [3, 15–23]. Therefore, in this thesis, the Bayesian methods introduced before are used to constrain the Wilson coefficients of dimension-six operators affecting the top quark.

The dimension-six operators considered in the following studies are introduced here and are given in the Warsaw basis [133]. Only operators affecting third-generation quarks are taken into account. Six operators modifying the interactions of the top quark with the bosons are considered:

$$\begin{aligned}
O_{uB} &= (\bar{q}_L \sigma^{\mu\nu} u_R)\, \tilde{\varphi} B_{\mu\nu}\,, & O_{uG} &= \left(\bar{q}_L \sigma^{\mu\nu} T^A u_R\right) \tilde{\varphi} G^A_{\mu\nu}\,, \\
O_{uW} &= \left(\bar{q}_L \sigma^{\mu\nu} \tau^I u_R\right) \tilde{\varphi} W^I_{\mu\nu}\,, & O^{(1)}_{\varphi q} &= \left(\varphi^\dagger i \overleftrightarrow{D}_\mu \varphi\right) (\bar{q}_L \gamma^\mu q_L)\,, \\
O_{\varphi u} &= \left(\varphi^\dagger i \overleftrightarrow{D}_\mu \varphi\right) (\bar{u}_R \gamma^\mu u_R)\,, & O^{(3)}_{\varphi q} &= \left(\varphi^\dagger i \overleftrightarrow{D}^I_\mu \varphi\right) \left(\bar{q}_L \tau^I \gamma^\mu q_L\right)\,.
\end{aligned} \tag{4.9}$$

In addition, five semileptonic four-fermion operators are included:

$$\begin{aligned}
O^{(1)}_{lq} &= \left(\bar{l}_L \gamma_\mu l_L\right) (\bar{q}_L \gamma^\mu q_L)\,, & O^{(3)}_{lq} &= \left(\bar{l}_L \gamma_\mu \tau^I l_L\right) \left(\bar{q}_L \gamma^\mu \tau^I q_L\right)\,, \\
O_{qe} &= (\bar{q}_L \gamma_\mu q_L) (\bar{e}_R \gamma^\mu e_R)\,, & O_{eu} &= (\bar{e}_R \gamma_\mu e_R) (\bar{u}_R \gamma^\mu u_R)\,, \\
O_{lu} &= \left(\bar{l}_L \gamma_\mu l_L\right) (\bar{u}_R \gamma^\mu u_R)\,.
\end{aligned} \tag{4.10}$$

For the studies of t-channel single top quark production cross sections in Chap. 6, also the four-quark operators

$$O_{qq}^{(1)} = (\bar{q}_L \gamma_\mu q_L)\,(\bar{q}_L \gamma^\mu q_L) \quad \text{and} \quad O_{qq}^{(3)} = \left(\bar{q}_L \gamma_\mu \tau^I q_L\right)\left(\bar{q}_L \gamma^\mu \tau^I q_L\right) \tag{4.11}$$

are considered. All of the operators above are constructed from the SM fields of the Lagrangian in Eq. (4.1). Therefore, q_L and l_L are again the $\mathrm{SU}(2)_L$ doublets of the quarks and leptons, u_R and e_R are the $\mathrm{SU}(2)_L$ singlets for the up-type quarks and charged leptons. The $B_{\mu\nu}$, $W^I_{\mu\nu}$, and $G^A_{\mu\nu}$ are the field strength tensors of $\mathrm{U}(1)_Y$, $\mathrm{SU}(2)_L$, and $\mathrm{SU}(3)_C$, respectively, and φ denotes the Higgs doublet. The σ^I are the Pauli matrices with $\tau^I = \sigma^I/2$. The Dirac gamma matrices are denoted as γ_μ, and T^A are the Gell-Mann matrices.

The sensitivities of the considered top-quark and B physics observables to these operators are discussed in the respective studies in Chap. 6, Chap. 7, and Chap. 8.

5 Combining measurements for EFT interpretations with EFTfitter.jl

The interpretation of measurements in terms of SMEFT is a popular approach for indirectly searching for BSM physics, as already discussed in the previous paragraphs. In this chapter, the need for a collective EFT interpretation of multiple measurements and different observables is indicated, and a Bayesian approach for combining correlated measurements is introduced. Current efforts targeting global SMEFT fits are motivated in Sec. 5.1, and statistical methods for combining correlated measurements of a single or multiple observables are discussed in Sec. 5.2. In Sec. 5.3, a new tool facilitating Bayesian combinations of measurements for EFT interpretations is introduced, and important aspects of its implementation are presented. Methods for determining the influence of individual measurements on the total combination are discussed in Sec. 5.4, and studies comparing the different weighting criteria are performed using a toy example.

5.1 Towards a global SMEFT fit in the top-quark sector

Since direct searches at the *Large Hadron Collider* (LHC) did not yield any hints for BSM physics, indirect searches using EFT approaches gained popularity during recent years. As motivated in Sec. 4.3.3, particularly LHC measurements of the top-quark sector are now frequently interpreted in terms of EFT, for example, by the ATLAS and CMS collaborations [8–14]. By comparing the measured values with corresponding SM predictions and translating potential deviations into contributions from dimension-six operators, constraints on the Wilson coefficients are derived. To improve the precision of the constraints, combining multiple measurements of the same observable is a common approach. As mentioned before, typically, only a certain number of dimension-six operators give non-negligible contributions to a specific process. Therefore, measurements of a single observable are usually sensitive to only a small subset of SMEFT coefficients. Combinations of different observables are pursued to provide higher sensitivity and to constrain larger sets of Wilson coefficients in a consistent way. Efforts for such combinations are ongoing both within the experimental collaborations and external working groups, targeting multi-observable and multi-coefficient SMEFT fits [5, 16, 17, 23, 136, 137]. LHC-wide working groups have formed to harmonize such efforts and to establish common standards for global SMEFT fits, for example, in the top-quark sector [15].

One of the challenges when aiming for a global fit of SMEFT Wilson coefficients is to find a suitable method that allows for a consistent combination of multiple observables, measured using different experiments, energy scales, and analysis methods. The most fundamental approach to combining measurements for a common EFT interpretation is to determine a combined likelihood function including all data available. However, such a likelihood-level combination of different experiments is an elaborate task in particle physics. A great amount of data has to be considered, and typically different techniques and various software frameworks are used to

derive the individual likelihoods. In particular, the models describing systematic uncertainties pose a challenge, as they usually cannot be merged in a simple way. Therefore, combining the likelihoods of measurements from different experiments in a coherent way is a sophisticated task that requires a lot of cooperation and harmonization of tools. Such combined analyses of ATLAS and CMS data have, for example, been performed in the Higgs sector [138, 139]. When combining different measurements and observables to interpret them in the context of SMEFT, usually simplified approaches are employed [2, 23, 140]. These methods then allow performing combined fits even outside the respective experimental collaborations and without access to the underlying data and likelihoods. In the following section, such approaches for the combination of correlated measurements are introduced.

5.2 Methods for combining measurements

The naive approach to combine multiple estimates of the same quantity is to determine the average. In the presence of uncertainties, the weighted average is often applied, with the weights being proportional to the inverse variances of the individual measurements. This approach, however, is only valid when the measurements are uncorrelated. In particle physics, correlations between different measurements can exist and then need to be considered. For the case of SMEFT fits of top-quark observables, this is also discussed in more detail in Chap. 6.

5.2.1 Correlated measurements

The results of high energy physics measurements are usually treated as normally distributed random variables, specified by a nominal value and an uncertainty. The uncertainty σ_i of a measurement with nominal value $\bar{x}_i$ is the square root of its variance $\sigma_i = \sqrt{\mathrm{Var}(x_i)}$, with the variance being the expected quadratic deviation from the expectation value, $\mathrm{Var}(x_i) = \mathrm{E}[(x_i - \bar{x}_i)^2]$. The joint variance of two measurements (of the same or of different quantities) is expressed by the covariance $\mathrm{cov}(x_i, x_j) = \mathrm{cov}(x_j, x_i) = \mathrm{E}[(x_i - \bar{x}_i) \cdot (x_j - \bar{x}_j)]$. The dependence between two measurements is often formulated in terms of the dimensionless linear correlation coefficient ρ, which is obtained by normalizing the covariance with the respective uncertainties

$$\rho_{ij} = \rho_{ji} = \frac{\mathrm{cov}(x_i, x_j)}{\sigma_i \sigma_j}\,, \tag{5.1}$$

and can take values $-1 \leq \rho_{ij} \leq 1$. In the case of two measurements being truly independent, the covariance and the correlation coefficient are zero. Conversely, however, the covariance and correlation coefficient being zero does not imply the independence of the measurements. For multiple measurements, the pairwise correlation coefficients or covariances are summarized in

symmetric and positive semi-definite correlation or covariance matrices,

$$M_{\text{cor}} = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \dots \\ \rho_{21} & 1 & \rho_{23} & \dots \\ \rho_{31} & \rho_{32} & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad \text{or} \quad M_{\text{cov}} = \begin{pmatrix} \text{var}(x_1) & \text{cov}(x_1, x_2) & \text{cov}(x_1, x_3) & \dots \\ \text{cov}(x_2, x_1) & \text{var}(x_2) & \text{cov}(x_2, x_3) & \dots \\ \text{cov}(x_3, x_1) & \text{cov}(x_3, x_2) & \text{var}(x_3) & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{5.2}$$

with the diagonal of the correlation matrix being unity by definition (see Eq. (5.1)) and the diagonal elements of the covariance matrix being the variances of the individual measurements.

Approaches for combining a set of measurements for which these matrices are not diagonal, i.e., where the correlation matrix is not the identity matrix, are introduced in the following paragraphs.

5.2.2 Combining correlated measurements

... of a single observable When combining multiple measurements of the same quantity, a common approach to account for correlations between the measurements is to determine the so-called *Best Linear Unbiased Estimator* (BLUE) [141]. As the name implies, the resulting estimate has the three properties of being (1) a linear combination of the input measurements, (2) an unbiased estimate, and (3) the best estimate, in the sense that it has the minimum possible variance. When combining N different measurements $\{x_i\}$ (with $i = 1, \dots, N$) of the same observable, the BLUE and its variance are given as

$$\hat{x} = \sum_{i=1}^{N} w_i x_i \quad \text{and} \quad \sigma_{\hat{x}}^2 = w^\top M w\,, \quad \text{where} \quad w_i = \left[\frac{M^{-1}U}{U^\top M^{-1} U} \right]_i . \tag{5.3}$$

Here, the matrix M is the covariance matrix of the measurements, and U is a vector with N elements all being unity. The elements w_i are the weights of the individual measurements in the linear combination. They are determined to minimize the variance of the resulting estimate and fulfill the normalization condition $\sum_i^N w_i = 1$. A feature of the BLUE method is that the covariance matrix of the input measurements can be composed of different sources of uncertainties, such as statistical and systematic uncertainties, for example,

$$M_{ij} = \text{cov}(x_i, x_j) = \sum_{k=1}^{K} \text{cov}^{(k)}(x_i, x_j)\,, \tag{5.4}$$

where k iterates over the K different types of uncertainties.

... of different observables When combining measurements of different observables, performing the BLUE combination described above separately for each observable only yields the best estimate if no correlations exist between measurements of the different observables. To account for such correlations, Eq. (5.3) can be generalized for correlated measurements of

different quantities [142]. For N measurements $\{x_i\}$ (with $i = 1, \ldots, N$) of n different observables y_α (with $\alpha = 1, \ldots, n$), where each observable y_α is measured at least once, i.e., $N_\alpha \geq 1$ and $N = \sum_\alpha N_\alpha \geq n$, the BLUE estimates and the corresponding covariance matrix of the estimates are given as

$$\hat{y}_\alpha = \sum_{i=1}^{n} w_{\alpha i} y_i \quad \text{and} \quad \text{cov}(\hat{y}_\alpha, \hat{y}_\beta) = w M w^\top , \quad \text{where} \quad w_{\alpha i} = \left[\frac{U^\top M^{-1}}{U^\top M^{-1} U} \right]_{\alpha i} . \tag{5.5}$$

Here, U is a matrix with the elements being

$$U_{i\alpha} = \begin{cases} 1, & \text{if } x_i \text{ is a measurement of } y_\alpha \\ 0, & \text{else} . \end{cases} \tag{5.6}$$

An estimate of the BLUE value that is equivalent to that of Eq. (5.5) can be determined by finding the values $\hat{y}_\alpha$ that minimize the weighted sum of squares

$$S = \sum_i \sum_j [\vec{x} - U\vec{y}]_i M_{ij}^{-1} [\vec{x} - U\vec{y}]_j . \tag{5.7}$$

This sum of squares in Eq. (5.7) can also be interpreted as a likelihood function L, using the relation

$$S = -2 \ln L , \tag{5.8}$$

of Wilks' theorem [143]. This likelihood then allows performing combinations of correlated measurements using Bayesian inference. When choosing uniform prior distributions, the BLUE estimate is recovered as the mode of the posterior distribution, with the variance (covariance matrix in the case of multiple observables) corresponding to the variance (covariance) of the posterior distribution.

... for inference on underlying model parameters In many cases, such as in EFT fits, combinations of measurements are not only performed to obtain the best combined estimate but rather to gain knowledge about common model parameters that are not directly observable but can be inferred from the measurements. In the case of SMEFT interpretations, for example, measurements of observables, such as cross sections and branching ratios, are combined in order to constrain Wilson coefficients as the underlying model parameters. In such cases, the Bayesian approach to the combination, i.e., using the likelihood of Eq. (5.8), can be generalized by introducing an uncertainty propagation step [2]. For this purpose, the observables $\vec{y}$ are assumed to depend on the free parameters $\vec{\lambda}$ of the model, i.e., $\vec{y} = \vec{y}(\vec{\lambda})$. The likelihood thus becomes a function of the measurements $\vec{x}$, given the model parameters $\vec{\lambda}$,

$$\ln p(\vec{x}|\vec{\lambda}) = -\frac{1}{2} \sum_i \sum_j [\vec{x} - U\vec{y}(\vec{\lambda})]_i M_{ij}^{-1} [\vec{x} - U\vec{y}(\vec{\lambda})]_j . \tag{5.9}$$

This likelihood can be interpreted as a measure of the distance between the observed values $\vec{x}$ and the predicted values for a given configuration of the model parameters $\vec{y}(\vec{\lambda})$. When performing Bayesian inference with this likelihood, the corresponding prior is a distribution of the model parameters $\vec{\lambda}$.

5.3 EFTfitter.jl

A tool that facilitates the application of the Bayesian approach to the combination of measurements based on Eq. (5.9) is EFTfitter [2]. While this approach and the corresponding likelihood are generally applicable to all use cases where the common assumption of Gaussian uncertainties holds, EFTfitter is specifically intended for EFT interpretations of high energy physics measurements.

The main objective of EFTfitter is to provide an interface that facilitates a user-friendly implementation of the likelihood in Eq. (5.9). For this purpose, in EFTfitter, the likelihood is automatically derived from user inputs for the underlying model, measurements, uncertainties, and correlations. The central element when applying EFTfitter is the user-specified model $\vec{y}(\vec{\lambda})$. It must be provided in terms of functions that return predictions for the values of the observables and depend only on the free model parameters to be constrained. In these functions, the underlying physics model, such as an EFT, enters the analysis. For SMEFT interpretations, this typically means that the dependences of observables, such as cross sections, on the Wilson coefficients have to be expressed in terms of simple functions. Furthermore, measurements of the observables have to be provided. As in the approach of Eq. (5.9) the measurements are assumed to have Gaussian uncertainties, they are specified as nominal values with corresponding values of the uncertainties. For the uncertainties, a breakdown of different sources, such as statistical or systematic uncertainties, for example, can be provided. For each type of uncertainty, an associated correlation matrix has to be specified. The total covariance matrix is automatically determined according to Eq. (5.4) as the sum over the individual covariance matrices.

For performing the Bayesian inference, EFTfitter relies on BAT. The first version of EFTfitter was released in 2016 and was implemented in C++, depending on the C++ version of BAT [2].[1] To take full advantage of the additional features and novel algorithms of BAT.jl, a new version of EFTfitter, called EFTfitter.jl[2] has been implemented in Julia within the scope of this thesis. The general workflow for using EFTfitter.jl is similar to that of the previous version. However, several new features have been implemented, improving its capabilities.

As mentioned above, the central task for the user is to provide an underlying physics model by specifying the dependences of the observables on the model parameters. In EFTfitter.jl, this is done by implementing plain Julia functions depending only on the model parameters to be constrained.

[1] `https://github.com/tudo-physik-e4/EFTfitterRelease`
[2] `https://github.com/tudo-physik-e4/EFTfitter.jl`

```
function observable1(parameters)
    return 2*parameters.λ1^2 + 0.4*parameters.λ1*parameters.λ2
end
```

Using ordinary Julia functions provides full flexibility in implementing the underlying model. Thus, it can also be derived from external tools or depend on interfaces to other languages, for example. An important aspect is that the functions returning the predictions of the observables need to be computationally fast since they are evaluated in each step of the sampling process. For SMEFT fits, this usually means that the rather slow MC simulations, which are typically used to determine BSM contributions to the observables, first need to be approximated to obtain faster parametrizations. The values of the measurements are entered as a Julia `NamedTuple`, consisting of EFTfitter.jl's `Measurement` objects. A name is assigned to each measurement, and the associated observable is specified. The measured nominal value is inserted, together with values for (multiple types of) uncertainties.

```
measurements = (
  Meas1 = Measurement(observable1, 21.6, uncertainties=(stat=0.8, syst=1.8),
  ... )
```

Correlations between the measurements are specified by providing a correlation matrix (or a `NoCorrelation` object for the identity matrix) for each type of uncertainty.

```
correlations = (
  stat = NoCorrelation(),
  syst = Correlation([1.0 0.5; 0.5 1.0]),
  ... )
```

After providing a BAT.jl prior distribution for the model parameters, an `EFTfitterModel` can be constructed, which automatically implements the likelihood of Eq. (5.9) from the user inputs.

```
model = EFTfitterModel(parameters, measurements, correlations)
posterior = PosteriorDensity(model)
```

Bayesian inference on the model parameters is then performed using BAT.jl and the algorithms discussed in Sec. 3.3 for sampling the corresponding posterior distribution. Compared to its predecessor version, the new implementation of EFTfitter.jl simplifies the use of distributions. While each bin of a distribution is counted as an individual measurement in the likelihood of Eq. (5.9), distributions can now be inserted as a joint `MeasurementDistribution` object. This facilitates the use of binned observables, such as differential cross sections.

```
MeasDist = MeasurementDistribution(diff_xsec, [1.9, 2.9, 4.4],
           uncertainties = (stat=[0.7, 1.1, 1.2], syst=[0.6, 0.7, 0.9]),
           active = [true, false, true])
```

When needed, also access to the individual bins is possible. As for all `Measurement` objects, also for the distributions, the keyword argument `active` can be used to exclude individual measurements or bins from the combination. Furthermore, the input of correlations is simplified in EFTfitter.jl. While in the predecessor version the full correlation matrices for each uncertainty had to be provided explicitly, in EFTfitter.jl, the correlation matrices can be constructed from the names of the measurements and corresponding correlation coefficients using the function `to_correlation_matrix`.

```
corr_matrix = to_correlation_matrix(measurements,
                (:Meas1, :Meas2, 0.4),
                ... )
```

The full correlation matrix is automatically derived and converted into the corresponding covariance matrix when instantiating the `EFTfitterModel`. This syntax also supports the use of block matrices, e.g., for the bins of distributions. Like in the C++ version, unknown correlation coefficients can be treated as nuisance parameters in the fit. In such cases, for each of the unknown correlation coefficients, an additional parameter is introduced into the fit using a `NuisanceCorrelation` object. During the inference, the corresponding elements of the covariance matrix in Eq. (5.9) are then varied according to their prior distributions.

```
nuisance_correlations = (
    rho  = NuisanceCorrelation(:syst, :Meas1, :Meas2, Normal(0.3, 0.05)),
)
```

Since the total number of correlation coefficients scales quadratically with the number of measurements included, this approach is only suitable for a small number of unknown correlations compared to the number of measurements. Otherwise, the number of free parameters becomes too large to be constrained by the data available.

Julia plot recipes are implemented, which allow visualizing the EFTfitter.jl types, such as the `Measurement` and `Observable` objects, for specific configurations of the model parameters. Examples of such plots are shown in Fig. 6.2 of Chap. 6. Like its predecessor version, EFTfitter.jl includes methods to estimate the influence of individual measurements (or types of uncertainty) on the result of the combination. Based on the relative increase (decrease) in the size of the posterior distribution when a single measurement (uncertainty type) is excluded from the combination, weights for the individual measurements (uncertainty types) are estimated.

A more detailed discussion of this approach is given in the next section. In particular, a comparison of the EFTfitter approach with other approaches for determining the relative weight of a single measurement in the combination is performed. A multi-stage artificial example is used to demonstrate the difficulties of finding a generally valid approach for determining such weights.

5.4 Estimating the impact of individual measurements on their combination

When combining multiple measurements with the approaches described above, it is often of interest to determine how much each individual measurement affects the result of the combination. Different criteria that allow assigning weights to the individual measurements are introduced below and compared using a generic example.

Central value weights When using the BLUE method to combine multiple measurements of the same observable, the weights w_i in Eq. (5.3) can be used to estimate the relative impact of the individual measurements on their linear combination. These weights are further referred to as *central value weights* (CVWs)[3], following the notation of Ref. [144]. It should be noted that the central value weights can take negative values, which then indicates the presence of a high-correlation regime [144].

Relative importance weights As the negative values of the central value weights may lead to problems in their interpretation, the *relative importance weight* (RIW) of a single measurement,

$$\mathrm{RIW}_i = \frac{|w_i|}{\sum_{j=1}^{N} |w_j|}\,, \tag{5.10}$$

is sometimes used instead to circumvent this problem by using absolute values. This approach has been applied, for example, in BLUE combinations of top-quark mass measurements from Tevatron and LHC [145, 146]. However, the concept of the relative importance weights has been criticized as being an improper approach, and its limitations have been demonstrated in Ref. [144]. Instead, two so-called *information weights* have been proposed in Ref. [144], which are based on the *Fisher information* I.

Intrinsic information weights The *intrinsic information weight* (IIW) of a single measurement x_i is defined as

$$\mathrm{IIW}_i = \frac{1/\sigma_i^2}{1/I}, \quad \text{with} \quad I = \frac{1}{\sigma_{\hat{x}}^2}\,, \tag{5.11}$$

where $\sigma_{\hat{x}}^2$ is the standard deviation of the BLUE estimate, as given in Eq. (5.3). It is the ratio of the Fisher information of each individual measurement to the Fisher information of the total combination. In this approach, the influence of correlations between measurements is considered not to be assignable to a single measurement and is accounted for in an additional

[3] For improved readability, in this section, only the weight types that are also used in further chapters of this thesis are abbreviated in the text. In the labels of plots, the abbreviations are used for all types of weights.

weight

$$\mathrm{IIW}_{\mathrm{cor}} = \frac{I - \sum_{i=1}^{N} 1/\sigma_i^2}{I}\,, \tag{5.12}$$

fulfilling the normalization $\mathrm{IIW}_{\mathrm{cor}} + \sum_{i=1}^{N} \mathrm{IIW}_i = 1$ by definition [144].

Marginal information weights The *marginal information weight* (MIW) of a measurement is defined as the ratio of its marginal information to the total information in the combination. It is determined by performing the combination again, excluding the respective measurement,

$$\mathrm{MIW}_i = \frac{\Delta I_i}{I_{(N \text{ meas.})}} = \frac{I_{(N \text{ meas.})} - I_{(N-1 \text{ meas., i.e. all meas. except } i)}}{I_{(N \text{ meas.})}}\,. \tag{5.13}$$

A more detailed discussion of the information weights and a comparison with the central value weights and the relative importance weights is given in Ref. [144].

Measurement exclusion weights When using the Bayesian approach of Eq. (5.9) for combining measurements, the dependences of the observables on the model parameters have to be taken into account for estimating the influence of the individual measurements. However, these dependences can lead to the posterior distributions not having multivariate Gaussian shapes, as is generally the case for the simple BLUE combinations without inference on the underlying parameters. Therefore, in EFTfitter.jl an approach similar to that of the marginal information weights is used, which is also valid for non-Gaussian posteriors. To determine weights for N individual measurements, the combination is performed N times, each time excluding one of the measurements. This results in N different posterior distributions, each accounting for only $N-1$ of the measurements. Certain properties of the obtained posterior distributions are then compared to the posterior distribution of the full combination to estimate the influence of each individual measurement. The weights obtained from this approach are further referred to as *measurement exclusion weights* (MEWs). The properties of the posterior distributions compared in EFTfitter.jl to determine the MEWs are the sizes of the *highest density regions* (HDRs). This approach is motivated by the idea that the sizes of the HDRs containing a certain fraction of the total posterior probability, e.g., the HDR containing 90 % of the posterior probability, can be interpreted as a measure of the overall uncertainty of the combination result. For a one-dimensional posterior distribution, the size of the HDR is the total width of the smallest intervals containing the specified probability. In the two-dimensional case, it is the corresponding area, and in higher dimensions, it is the corresponding (hyper-)volume. Therefore, the term *highest density volume* (HDV) is used here to specify the size of a HDR. When referring explicitly to the one-dimensional and two-dimensional cases, the terms *highest density length* (HDL) or *highest density area* (HDA), respectively, are used for more precise phrasing.

By default, the 90 % HDVs are used in EFTfitter.jl to obtain estimates of the total uncertainty, as they are also valid for non-Gaussian or multi-modal posterior distributions. In such cases,

other measures, like the variance or covariance, are not necessarily meaningful. When removing a single measurement from the fit, the HDV is expected to increase because less information is available to constrain the parameter space. This is visualized in Fig. 5.1, where two-dimensional 90 % HDRs are shown, obtained by combining multiple measurements of a generic example. It is visible that when excluding individual measurements from the combination, the sizes of the corresponding HDRs increase compared to that of the total combination of all measurements. The corresponding values of the HDVs for the different posterior distributions are also shown in Fig. 5.1.

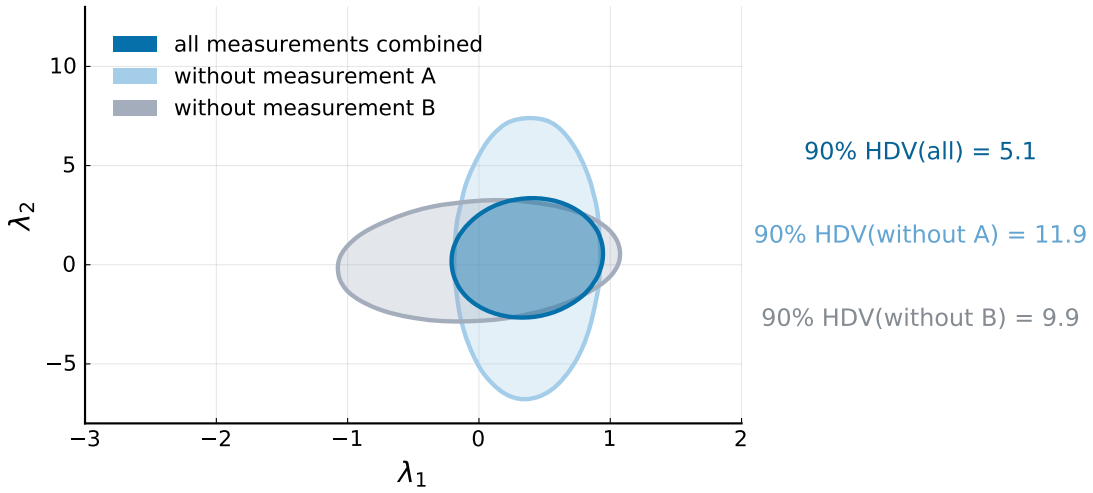

Figure 5.1: Visualization of the increase of HDVs when excluding individual measurements from the combination. Shown are the HDRs containing 90 % of the posterior probability of a generic example with two free parameters.

It should be noted, however, that excluding a single measurement from the combination can also result in a decrease of the HDV, e.g., in cases where one measurement contradicts all others. In EFTfitter.jl, the relative change of the HDV when excluding a single measurement is used to estimate its influence on the total combination. By removing one measurement at a time from the fit, the relative change of the HDV can be calculated for each measurement as

$$\delta\mathrm{HDV}_i = \frac{\mathrm{HDV}_i - \mathrm{HDV}_N}{\mathrm{HDV}_N}\,, \tag{5.14}$$

where HDV_i is the HDV when excluding measurement x_i, and HDV_N is the HDV of the total combination including all N measurements. The corresponding MEWs are then obtained by normalizing according to

$$\mathrm{MEW}_i = \frac{\delta\mathrm{HDV}_i}{\sum_i^N \delta\mathrm{HDV}_i}\,. \tag{5.15}$$

In particular, it should be mentioned that this approach can be applied not only to estimate the impact of individual measurements but also the impact of different types of uncertainties. For this purpose, the combination can be performed several times, each time excluding a certain uncertainty category. Since fewer uncertainties are accounted for in the individual fits,

the corresponding posterior distributions are expected to be narrower. The resulting HDVs are thus expected to decrease compared to the fit where all types of uncertainties are considered. Therefore, in EFTfitter.jl, the relative decrease of the HDVs can be used to assign weights to the different uncertainty categories, the so-called *uncertainty exclusion weights* (UEWs).

Estimating the HDR in EFTfitter.jl

In EFTfitter.jl, the HDR containing a certain amount of probability is determined numerically from binned posterior samples. The weighted sum over the bins is computed starting from the bin with the largest weight until the summed probability has reached the desired threshold. By multiplying the number of bins belonging to the HDR with the volume of each bin, the size of the HDR, i.e., the HDV, is obtained. When combining measurements of a single observable without underlying model parameters and using a uniform prior, i.e., in a BLUE-like scenario, the posterior obtained with EFTfitter.jl is a one-dimensional normal distribution. In such cases, the size of the HDR is the length of the central interval containing the requested probability. In particular, the size of the 68.3 % HDR then corresponds to the variance $\sigma^2_{\hat{x}}$ of the normal distribution. Thus, in such cases, also the variance estimated from the posterior samples can be used in EFTfitter.jl for determining MEWs, which is equivalent to using the 68.3 % HDV.

5.4.1 Example: Combining measurements of a single observable

The different criteria for estimating the influence of individual measurements described above are now compared. For this purpose, different scenarios of the artificial example on W boson decays, introduced in Sec. 4 of Ref. [142], are adopted here and extended in the following.

Suppose two experiments, A and B, measure the branching fraction $\mathcal{B}$ of leptonic W boson decays separately in the electron, muon, and tau channels. The measured nominal values and uncertainties are given as

$$\begin{aligned} A_e &= (10.5 \pm 1.0)\,\% & B_e &= (13.5 \pm 3.0)\,\% \\ A_\mu &= (10.2 \pm 0.9)\,\% & B_\mu &= (12.3 \pm 1.5)\,\% \\ A_\tau &= (9.50 \pm 3.0)\,\% & B_\tau &= (14.0 \pm 3.0)\,\% \ . \end{aligned}$$

The numbers of the electron and tau channels are taken from the example in Ref. [142], while the numbers for the muon channel are added here for the purpose of the following studies.

For a first comparison, these measurements are combined under the assumption of lepton flavor universality. Thus, all six measurements are assumed to describe the same underlying observable, i.e., the branching ratio $\mathcal{B}_\ell$ of a W boson decaying into a lepton. Throughout this example, the Bayesian combinations are always performed using uniform prior distributions. As mentioned above, in this case, the result of the EFTfitter.jl combination corresponds to the BLUE estimate, which in the absence of correlations yields $\hat{\mathcal{B}}_\ell = (10.85 \pm 0.58)\,\%$. The different types of weights are computed for this example. Specifically, these are the central

value weights, the relative importance weights, the intrinsic information weights, and the marginal information weights of the BLUE method, as well as the EFTfitter.jl MEWs using the HDRs containing 68.3 % and 95.5 % of the posterior probability. In addition, the MEWs are computed once using the variance estimate of the EFTfitter.jl combination, where the variance is estimated numerically from the posterior samples, and once using the analytically determined variance of the BLUE method.

In the top plot of Fig. 5.2, the resulting weights are shown for the lepton-universal combination without correlations between the measurements. It should be noted that, according to their definition in Eq. (5.13), the marginal information weights are not normalized. For better comparison with the other weights, always normalized weights are shown, obtained similarly to Eq. (5.15).

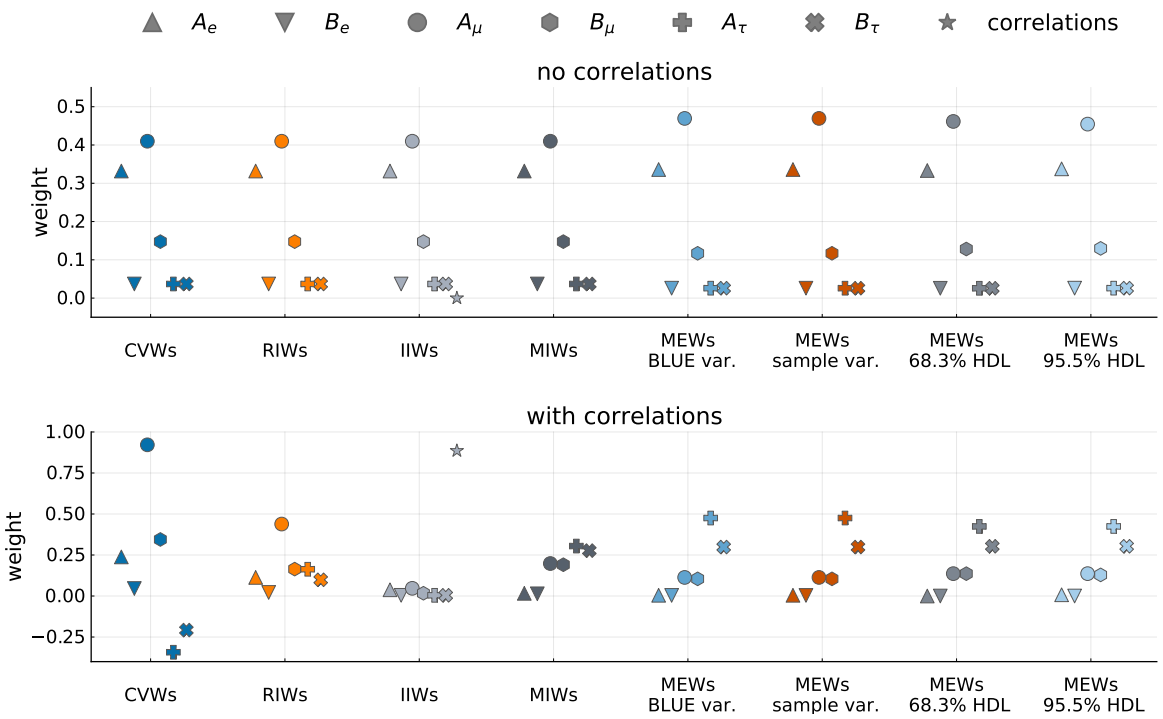

Figure 5.2: Comparison of the different types of weights for the lepton-universal example when assuming no correlations between the measurements (top) and when assuming the correlations of Fig. B.1 (bottom).

As can be seen in the upper plot of Fig. 5.2, in the uncorrelated case, all types of weights show a very similar hierarchy of the individual measurements, with A_μ and A_e having the largest influence, and the measurements B_e, A_τ, and B_τ having very small weights close to zero. This also agrees well with the naive expectations when considering only the magnitudes of the associated uncertainties of the measurements. In particular, the central value weights, the relative importance weights,the intrinsic information weights, and the marginal information weights have very similar values. Also, the MEWs all agree very well among each other. For the BLUE marginal information weights and BLUE variance MEWs, it should be noted

that they are based on the same information, with the only difference being the inversion of the variance in the case of the marginal information weights. As expected, contributions to the intrinsic information weights coming from correlations, i.e., $\mathrm{IIW}_{\mathrm{cor}}$, are zero in this uncorrelated example.

In the next step, correlations between the measurements are introduced in order to investigate their influence on the weights. It is assumed that measurements of the different lepton channels from experiment A are all correlated rather strongly among each other, with $\rho = 0.8$. For experiment B, the correlations between the different channels are assumed to be weaker, with $\rho = 0.5$. Negative correlations with a value of $\rho = -0.3$ are assumed to be present between the measurements of experiments A and B for the same lepton channels, to also include negative correlation coefficients in the example. The full correlation matrix is given in Fig. B.1 of App. B. When assuming these correlations, the BLUE combination of the measurements yields $\hat{\mathcal{B}}_\ell = (10.60 \pm 0.20)\,\%$. The different types of weights obtained in this correlated scenario are shown in the bottom plot of Fig. 5.2. It is visible that the central value weights for measurements A_τ and B_τ now take negative values, which is a consequence of the large correlations [144]. It is noticeable that for the intrinsic information weights, the largest weight is attributed to the correlations, and the weights for the individual measurements are all very close to zero. Again, the weights based on excluding a single measurement all agree very well and yield the same hierarchy of the measurements. Due to the presence of correlations, however, the measurements with the smallest uncertainties no longer receive the largest weights. In particular, for the marginal information weights, it can be noticed that the weights for A_τ and B_τ are very similar, whereas for all MEWs, A_τ obtains a noticeably larger weight than B_τ. It should be noted that the results in the bottom plot of Fig. 5.2 apply, of course, only for this one example of many possible correlation scenarios. However, in all other correlation scenarios tested, it was also observed that the different MEWs agree well among each other and yield similar hierarchies as the marginal information weights, with the latter usually being closer together, as seen in Fig. 5.2.

5.4.2 Example continued: Combining measurements of different observables

The above example of W boson decays is continued, but the assumption of lepton flavor universality is now dropped so that the three lepton channels are considered separate observables. Therefore, when combining the measurements from experiments A and B, three different estimates, $\hat{\mathcal{B}}_e$, $\hat{\mathcal{B}}_\mu$, and $\hat{\mathcal{B}}_\tau$, are obtained. In the uncorrelated scenario, the results of the BLUE combination are

$$\hat{\mathcal{B}}_e = (10.79 \pm 0.95)\,\%\,, \quad \hat{\mathcal{B}}_\mu = (10.76 \pm 0.77)\,\%\,, \quad \hat{\mathcal{B}}_\tau = (11.75 \pm 2.12)\,\%\,.$$

To estimate the influence of the individual measurements, the BLUE weights are calculated separately for each of the observables. In the Bayesian approach, this corresponds to calculating the MEWs based on the respective one-dimensional marginal distributions. These weights for the muon channel are shown in Fig. 5.3, again for both the uncorrelated and the correlated scenario. Corresponding plots for the electron and tau channel are given in Fig. B.2 and Fig. B.3 of App. B, respectively.

In the uncorrelated case, as expected, only the two measurements of the considered channel give contributions to the combination, and the weights of the other four measurements are zero. In the correlated case, also the weights of other measurements take values slightly above zero or, in the case of the central value weights, even slightly negative weights.

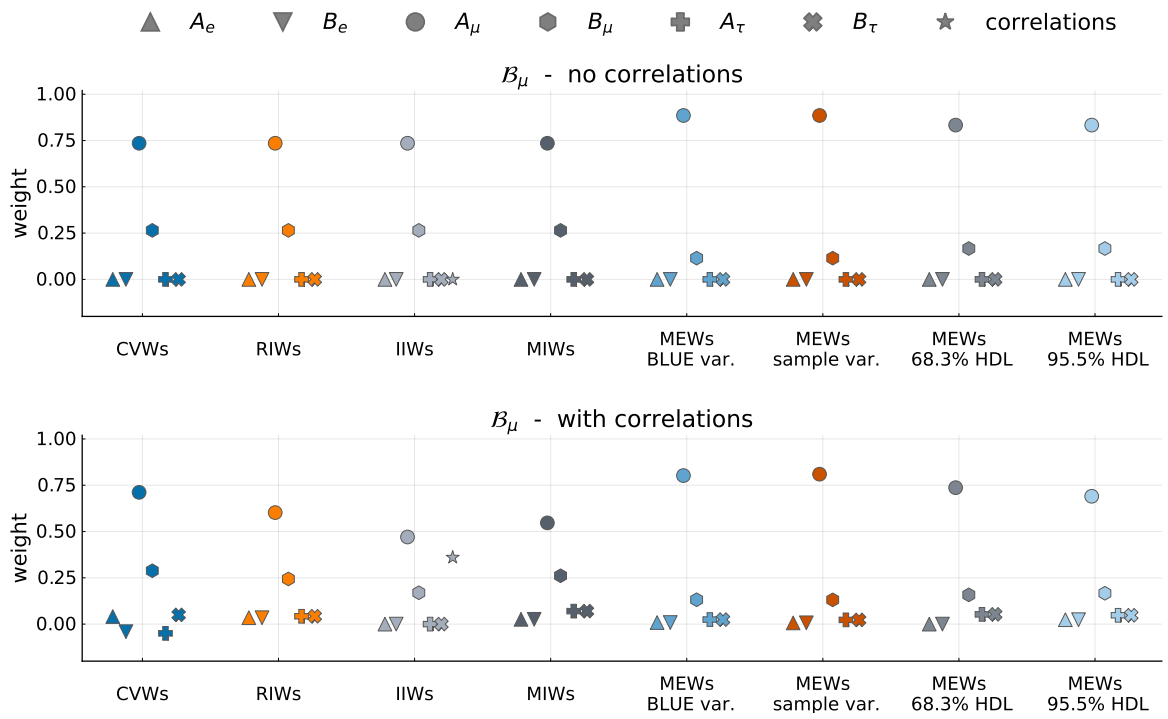

Figure 5.3: Comparison of the different types of weights for the muon channel when assuming no correlations between the measurements (top) and when assuming the correlations of Fig. B.1 (bottom).

In particular when using the Bayesian approach implemented in EFTfitter.jl, it might also be of interest to estimate which of the measurements has the largest overall impact on the (multidimensional) posterior distribution. In the current example, no inference on underlying model parameters is performed, but only the best combined estimate is determined. Thus, according to Eq. (5.8), always a multivariate normal posterior distribution is obtained. In this case, the generalized variance, i.e., the determinant of the covariance matrix of the combination results, can be used as an estimate of the overall uncertainty. In the more general approach of Eq. (5.9), the observables can have arbitrary dependence on underlying model parameters. As mentioned above, in such cases, the posterior distribution is not necessarily a multivariate normal distribution but can have complex shapes and multiple modes. Therefore, in EFTfitter.jl, typically, the size of a HDR of the posterior distribution is considered as a measure of the overall uncertainty, which is also applicable for non-Gaussian posterior shapes. Computing a HDV of the three-dimensional posterior distribution is, thus, another approach for calculating overall weights in this example. Estimating the HDV from the posterior samples, however, is a non-trivial task in multiple dimensions, as will be pointed out in the next paragraph.

Estimating the HDV in multiple dimensions The default criterion for ranking the influence of individual measurements in EFTfitter.jl is based on the HDV. It is determined numerically from binned posterior samples, as described above. This approach, however, suffers from the curse of dimensionality and is not feasible for more than three dimensions, as will be demonstrated here. In the case of one- and two-dimensional posterior distributions, computing the HDV from binned samples yields reliable results when using the typical sample sizes obtained from the BAT.jl samplers, i.e., $\mathcal{O}(10^6) - \mathcal{O}(10^7)$ samples. For three dimensions, however, the results of the HDV calculation already depend strongly on the number of posterior samples and the number of bins used for the calculation. This problem is visualized in Fig. 5.4, where estimates of the 68.3 %, 95.5 %, and 99.7 % HDVs are shown as a function of the number of bins per dimension for a two-dimensional (left plot) and three-dimensional (right plot) uncorrelated standard normal distribution. The HDV estimates are shown for 10^6 and 10^7 posterior samples. The dashed lines indicate the true values of the HDVs, calculated as the volumes of n-spheres with corresponding radii.

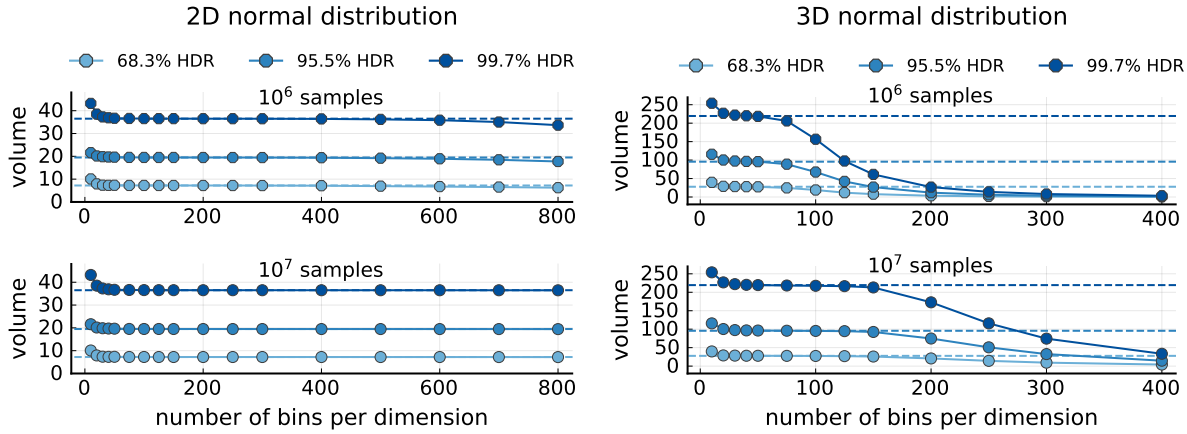

Figure 5.4: Estimated values of the 68.3 %, 95.5 %, and 99.7 % HDVs for a two-dimensional (left) and three-dimensional (right) standard normal distribution as a function of the number of bins used for the computation. The true values of the HDVs are indicated as dashed lines.

It is visible in Fig. 5.4 that in the two-dimensional case, the numerical HDV estimates are consistent for the typical number of bins per dimension (the EFTfitter.jl default is 200 bins per dimension) and yield estimates very close to the true value, both for 10^6 and 10^7 samples. Only for a very small number of bins, the HDV is overestimated, which is a result of an insufficient resolution due to the large bin size. In the three-dimensional case, a strong dependence on the number of bins is observed, and the true value of the HDV is only estimated correctly within a small range of the number of bins. The decreasing values of the HDV estimates for an increasing number of bins are a natural consequence of the ratio between the number of samples and the number of bins. If the phase space is divided into too many bins, in relation to the available number of samples, bins that would be part of the HDR are not populated by any posterior sample and are, thus, not included in the HDV computation. In the case of the three-dimensional normal distribution in the right plot of Fig. 5.4, for more than 100 bins per dimension, the total number of bins already exceeds the number of 10^6 samples available, and

thus the weights cannot be determined correctly. This leads to the observed underestimation of the HDVs. Since using much more than 10^7 samples is computationally unfeasible, different approaches for obtaining reliable estimates of the HDVs in higher dimensions are needed. While in the case of multivariate normal posterior distributions, the HDV can also be determined analytically, the posterior shapes obtained with EFTfitter.jl can be arbitrarily complicated. More sophisticated methods for estimating the HDV of the posterior distribution are needed to overcome the curse of dimensionality but are currently not available in the literature [147]. Therefore, only the naive approach based on binned samples is available in EFTfitter.jl so far. In particular, methods like nested sampling and AHMI might be suitable for developing algorithms to estimate HDVs in more than three dimensions.

Using marginal HDR estimates Because of these difficulties in obtaining the HDVs for higher-dimensional distributions, it is investigated if overall weights for the measurements can also be estimated using the one- and two-dimensional marginal distributions. In particular, it is studied if the summed HDLs and HDAs can sometimes be used to estimate MEWs that are compatible with the MEWs based on the full three-dimensional HDVs. Also, MEWs based on the generalized variance are included in the comparison. Overall weights from the BLUE weighting criteria are obtained by summing the weights of the individual observables and renormalizing them. Thus, the following criteria for determining overall weights are compared in the non-lepton-universal example:

- Renormalized sums over the BLUE weights for each observable
- MEWs based on the generalized variance, i.e., the determinant of the analytical (BLUE) covariance and the determinant of the numerical covariance estimated from the samples
- MEWs based on the determinants to the inverse power of the dimension
- MEWs based on the sum of all marginal 95.5 % HDLs and the MEWs based on the sum of all marginal 95.5 % HDAs
- MEWs based on the three-dimensional HDVs containing 95.5 % of posterior probability, once calculated analytically as the volume of an ellipsoid specified by the analytical covariance estimate of the BLUE method, and once estimated numerically from the posterior distribution using the binned method discussed above.

In Fig. 5.5, these overall weights for the measurements are shown for both the correlated and the uncorrelated example.

In the uncorrelated case, again, a good agreement is observed in the hierarchy of the different types of weights. In particular, the BLUE weights and the MEWs agree well among each other. That the MEWs based on the analytical and numerical estimation of the 95.5 % HDV agree very well, indicates that the number of bins for the numerical HDV estimation is well chosen. For the MEWs based on the generalized variance, the weight of A_e is estimated higher and that of B_μ lower than for the other MEWs. It is also noticeable that for this multivariate Gaussian posterior distribution, the MEWs based on the sums of the HDLs and HDAs yield

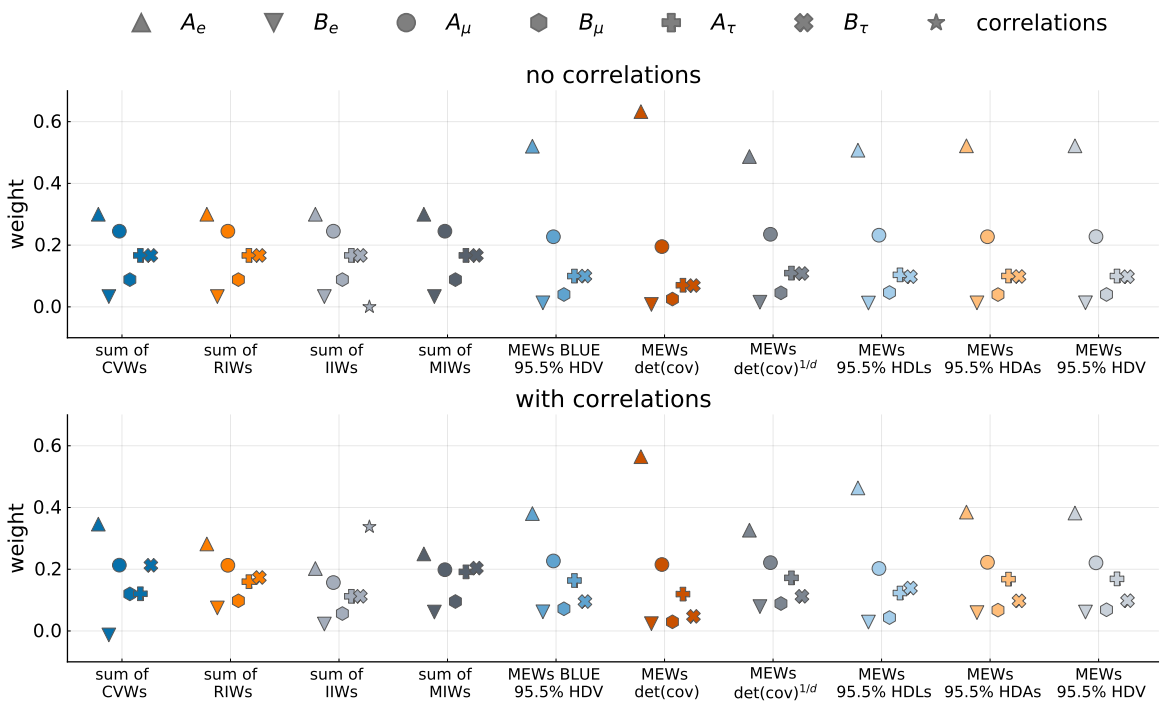

Figure 5.5: Comparison of the different types of overall weights for the non-lepton-universal example when assuming no correlations between the measurements (top) and when assuming the correlations of Fig. B.1 (bottom).

the same values as the MEWs coming from the three-dimensional HDVs. In the correlated case, the BLUE weights yield very different values and do not agree with the MEWs based on the HDVs, which again agree well among each other. The MEWs based on the generalized variance show the same hierarchy as those based on the HDVs, but different values of the specific weights, particularly for A_e. In this example, the weights based on the sum of all two-dimensional HDAs seem to be a good estimate for the weights based on the full three-dimensional HDV. However, the sum of the one-dimensional HDLs leads to different weights for A_τ and B_τ. This is not surprising since basically no information about correlations can be retrieved from the one-dimensional marginal distributions.

5.4.3 Example extended: Constraining underlying parameters

In the examples above, the combination of measurements has been performed without inference on underlying parameters. Thus, the resulting posterior distributions always had a (multivariate) Gaussian shape. As mentioned before, when performing combinations with EFTfitter.jl for inference on underlying parameters, this is not necessarily the case. The posterior distribution can take arbitrary shapes, depending on the relation between the observables and the parameters. In particular, this implies that the (generalized) variance of the posterior distribution is not always a reasonable quantity for determining weights. To investigate the compatibility of the different types of weights in such a typical EFTfitter.jl use case, a generic dependence of

the observables on fictitious underlying parameters λ_i is introduced. The specific dependence chosen for the purpose of the example is

$$\begin{aligned} \mathcal{B}_e(\lambda_1) &= 0.01(\lambda_1 - 5)^2 + 9.5 \\ \mathcal{B}_\mu(\lambda_1, \lambda_2) &= 0.01(\lambda_1 - 5)^2 + 7.5 + 0.01\lambda_2^2 \\ \mathcal{B}_\tau(\lambda_2, \lambda_3) &= \lambda_3 + 0.2\lambda_2 \,. \end{aligned} \tag{5.16}$$

The resulting posterior distribution when assuming a uniform prior is shown in Fig. 5.6 for the uncorrelated case. The corresponding posterior distribution when assuming the correlations of Fig. B.1 is shown in Fig. B.4 of App. B.

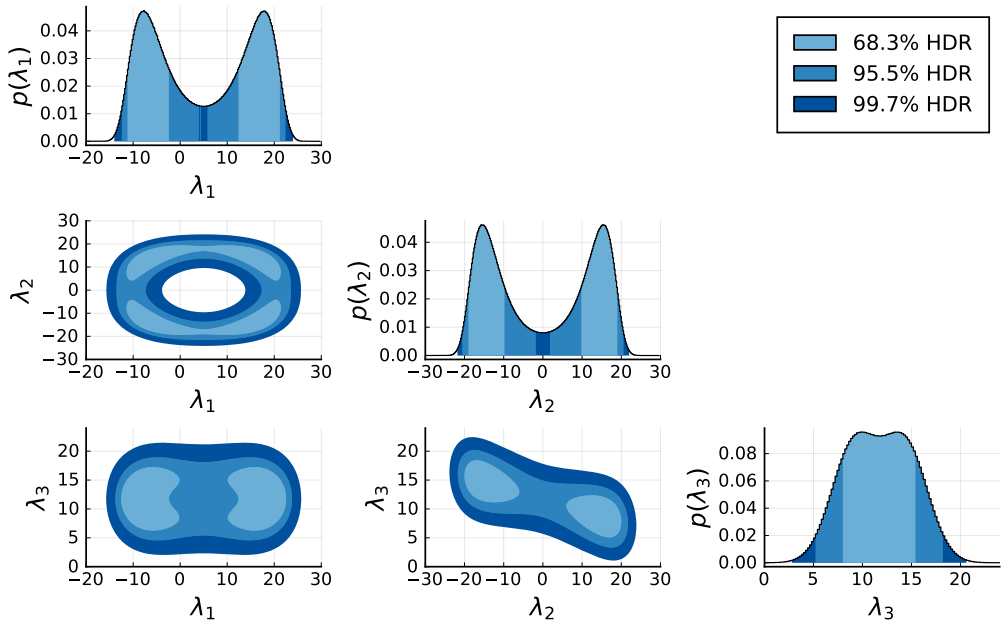

Figure 5.6: Posterior distribution of the total combination when assuming the dependence on underlying parameters of Eq. (5.16) in the uncorrelated case. The one- and two-dimensional marginal distributions are shown, and the HDRs containing 68.3 %, 95.5 %, and 99.7 % of the posterior probability are highlighted.

The marginalized posterior distributions of the parameters λ_1 and λ_2 are bimodal, and the associated two-dimensional marginal distribution has the shape of a ring. This shape of the distribution was chosen intentionally, as these are typical posterior distributions obtained in SMEFT fits with EFTfitter.jl, due to the often quadratic parametrizations of observables. Similar posterior shapes are also obtained in the SMEFT studies of Chap. 7.

For this example, only MEWs can be determined since the weights related to the BLUE method cannot be applied when inference on underlying parameters is performed. Therefore, again the MEWs based on the three-dimensional HDV are compared to MEWs based on the sums of the HDAs and HDLs. For this example, the 95.5 % and the 68.3 % HDRs are considered. To demonstrate that the generalized variance is not a universally reliable estimate of the total uncertainty, the MEWs based on the determinant of the covariance matrix are also included

in the comparison. The weights obtained in this example are given in Fig. 5.7, again for both the uncorrelated and the correlated case.

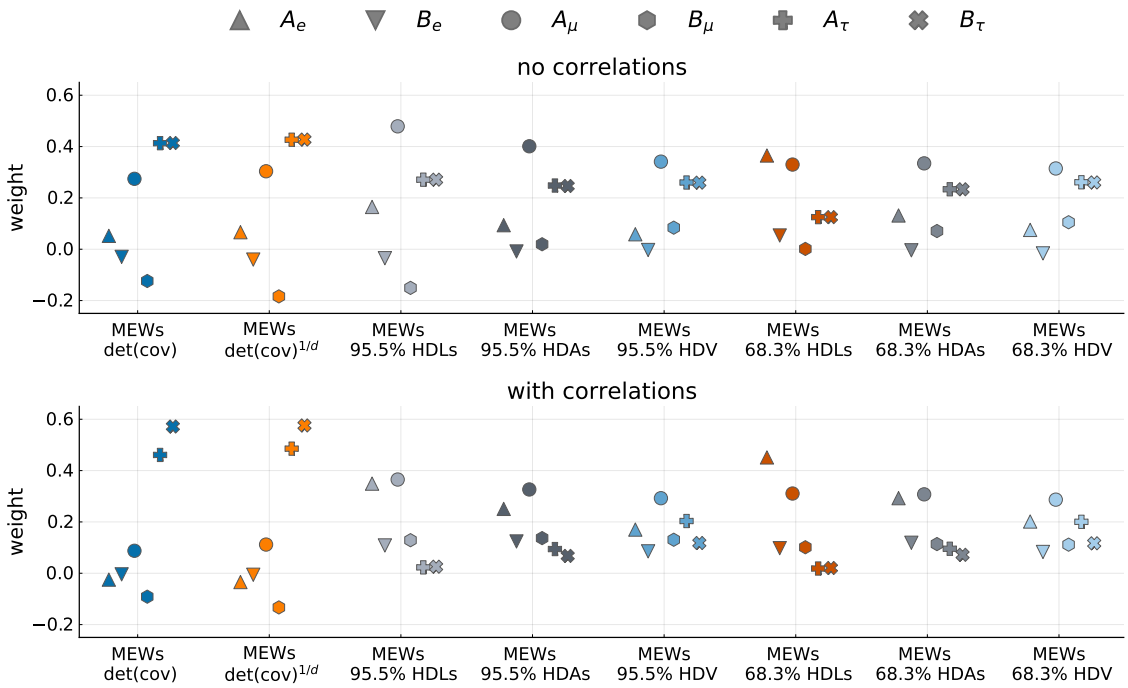

Figure 5.7: Comparison of the different types of overall weights for the example with inference on underlying parameters when assuming no correlations between the measurements (top) and when assuming the correlations of Fig. B.1 (bottom).

It can be observed in Fig. 5.7 that the weights based on the covariance matrix do not agree with any of the other weights. This is expected from the posterior distribution in Fig. 5.6 since the ring-shaped distribution of λ_1 and λ_2 cannot be described properly by the covariance matrix. In the uncorrelated case, the weights obtained from the 95.5 % and 68.3 % HDVs agree very well in their hierarchies, with only minor differences in the values for A_e and B_e. This is also observed in the correlated case. The MEWs obtained from the sums of HDLs differ for the 95.5 % and 68.3 % HDRs and do not agree with the HDV weights regarding the hierarchy of the measurements, in both correlation scenarios. In the uncorrelated case, the MEWs based on the sum of the HDAs appear to be rather compatible with the weights obtained from the HDVs at first glance. However, for both the 95.5 % and the 68.3 % HDRs, the weight of B_μ is lower than that of the HDVs, and thus, the hierarchy of the measurements is also different. In the correlated case, a similar effect can be observed, and the hierarchy of measurements based on the HDAs is different from that based on the HDVs, mainly due to differences in the weight of A_τ. Such differences are expected as the full three-dimensional posterior shape cannot be approximated using only the two-dimensional projections.

Summarizing these studies, it can be noted that MEWs based on HDRs provide a reasonable option for assigning weights to the individual measurements in their combination, even in

cases in which the posterior distributions have non-Gaussian shapes. In the simplest cases, the MEWs agree well with the weights obtained by the BLUE method. However, it was pointed out that for posterior distributions with more than three parameters, a reliable numerical estimation of the HDVs is currently not feasible. In the examples above, it has been demonstrated that an approximation using the one- or two-dimensional marginal distributions is possible in some cases but is not necessarily valid in general. Therefore, for multidimensional posterior distributions, it is currently impossible to obtain universally valid and reliable estimates for overall weights of the measurements. In such cases, it is rather recommended to determine the MEWs for the individual one- and two-dimensional marginal posterior distributions and to estimate the impact of the measurements based on these results.

6 Studying the impact of correlated measurements on the results of SMEFT fits

The studies in the previous chapter already indicated that correlations between measurements can have a sizeable impact on the results of their combination. However, since the values of correlation coefficients are not always available, measurements are often treated as being uncorrelated in such combinations. In this chapter, the potential impact of correlations on the results of SMEFT fits with EFTfitter.jl is examined systematically. For this purpose, constraints on three Wilson coefficients of dimension-six operators are derived using measurements of single top-quark production and top-quark decay processes, and the results of fits assuming different correlation scenarios are compared. This chapter is organized as follows: In Sec. 6.1, the common complications in the treatment of correlations in SMEFT fits are described. The EFT operators to be constrained are given in Sec. 6.2, and the observables and measurements included in the fits are introduced in Sec. 6.3. The parametrization of BSM contributions in terms of Wilson coefficients is discussed in Sec. 6.4. The uncertainty categories considered in the analysis, as well as the assumptions for the different correlation scenarios, are introduced in Sec. 6.5. The SMEFT fits and comparisons of the different correlation scenarios are performed in Sec. 6.6, and the impact of the correlations is studied in detail. In Sec. 6.7, an outlook is given on how the influence of correlations might evolve when future measurements with reduced uncertainties are taken into account. The findings of the studies in this chapter are briefly summarized in Sec. 6.8.

6.1 SMEFT fits with correlated measurements

As motivated in Sec. 5.1, fits combining measurements of different observables and from multiple experiments are commonly performed to constrain the Wilson coefficients of SMEFT operators. In such fits, potential correlations between uncertainties of the measurements are often not accounted for. This is because the values of respective correlation coefficients are only rarely available. Estimating correlations between measurements of different observables, particularly between measurements from different experiments, is a complex task and is virtually impossible for researchers outside the respective experimental collaborations. However, global SMEFT fits combining measurements from different experiments are often performed by working groups outside the experimental collaborations. As a consequence, the simplifying assumption of uncorrelated measurements is regularly used in these fits if no other information is publicly available [23, 136]. Another approach commonly applied to circumvent the problem of unknown correlations is to include only measurements for which no (or only very small) correlations are expected. This usually implies that only one measurement of each observable is considered, even if several measurements are available. For these reasons, the impact correlations can potentially have on the results of SMEFT fits is studied in the following. Special focus is placed on comparing the resulting constraints on the Wilson coefficients obtained from fits with uncorrelated and correlated uncertainties.

6.2 Wilson coefficients to be constrained

In the following, SMEFT Wilson coefficients of the top-quark sector are constrained using a combination of single top-quark production and top-quark decay measurements. The leading BSM contributions arise from the interference of dimension-six operators with the SM. Four of the SMEFT operators introduced in Sec. 4.3.3 contribute at order $\mathcal{O}(\Lambda^{-2})$ to these top-quark processes and are therefore considered here,

$$\begin{aligned} O_{\varphi q}^{(3)} &= \left(\varphi^\dagger i \overleftrightarrow{D}^I_\mu \varphi\right)\left(\bar{q}_L \tau^I \gamma^\mu q_L\right) , & O_{qq}^{(1)} &= \left(\bar{q}_L \gamma_\mu q_L\right)\left(\bar{q}_L \gamma^\mu q_L\right) , \\ O_{uW} &= \left(\bar{q}_L \sigma^{\mu\nu} \tau^I u_R\right)\tilde{\varphi} W^I_{\mu\nu} , & O_{qq}^{(3)} &= \left(\bar{q}_L \gamma_\mu \tau^I q_L\right)\left(\bar{q}_L \gamma^\mu \tau^I q_L\right) . \end{aligned} \tag{6.1}$$

Neglecting suppressed contributions proportional to masses much smaller than the top-quark mass, i.e., $m \ll m_t$, the leading BSM contributions to the observables considered in this analysis can be parametrized using only the three coefficients

$$\tilde{C}_{\varphi q}^{(3)} , \quad \tilde{C}_{uW} , \quad \tilde{C}_{qq} = \tilde{C}_{qq}^{(3)1133} + \frac{1}{6}\left(\tilde{C}_{qq}^{(1)1331} - \tilde{C}_{qq}^{(3)1331}\right) . \tag{6.2}$$

Here, additional superscripts denote the quark flavors, and $\tilde{C}_i$ are rescaled coefficients, defined as

$$\tilde{C}_i = \frac{v^2}{\Lambda^2} C_i , \tag{6.3}$$

where $v = 246\,\text{GeV}$ is the vacuum expectation value of the Higgs field. In the following, all Wilson coefficients are assumed to be real-valued, since only the coefficient of the non-Hermitian operator O_{uW} could take complex values, but just its real part contributes at $\mathcal{O}(\Lambda^{-2})$ to the considered processes.

6.3 Observables and measurements

This study includes measurements of total and differential single top-quark t-channel production cross sections, as well as measurements of top-quark decay observables, specifically the W boson helicity fractions and the top-quark decay width. Counting the bins of differential distributions as individual observables, a total of 58 measurements of 41 different observables is included. In the following, the observables and corresponding measurements considered are introduced. An overview of all observables and measurements is given in Tab. 6.1, where also the corresponding references are listed.

6.3.1 Single top-quark production cross sections

Singly-produced top quarks have first been observed by the CDF and D0 experiments at the Tevatron [166, 167]. Although the dominant production of top quarks at the LHC happens via pair production, precise measurements of single top quarks at different center-of-mass energies have been performed by the ATLAS and CMS experiments, allowing to study the properties

Table 6.1: Measurements of top-quark production and decay observables considered in the analyses. Also the respective center-of-mass energies ($\sqrt{s}$), the integrated luminosities, the experiments, the specific observables, and the corresponding publication references are listed.

Process	$\sqrt{s}$	Luminosity	Experiment	Observable	Reference
Single top	7 TeV	$4.59\,\text{fb}^{-1}$	ATLAS	$\sigma(tq)$, $\sigma(\bar{t}q)$, $\frac{d\sigma(tq)}{dp_T}$, $\frac{d\sigma(\bar{t}q)}{dp_T}$	[148]
		$1.17\,\text{fb}^{-1}(\mu)$	CMS	$\sigma(tq+\bar{t}q)$	[149]
		$1.56\,\text{fb}^{-1}$(e)	CMS	$\sigma(tq+\bar{t}q)$	[149]
Single top	8 TeV	$20.2\,\text{fb}^{-1}$	ATLAS	$\sigma(tq)$, $\sigma(\bar{t}q)$, $\frac{d\sigma(tq)}{dp_T}$, $\frac{d\sigma(\bar{t}q)}{dp_T}$	[150]
		$19.7\,\text{fb}^{-1}$	CMS	$\sigma(tq)$, $\sigma(\bar{t}q)$, $\sigma(tq+\bar{t}q)$, $\frac{d\sigma}{d\lvert y(t/\bar{t})\rvert}$	[151, 152]
Single top	13 TeV	$3.2\,\text{fb}^{-1}$	ATLAS	$\sigma(tq)$, $\sigma(\bar{t}q)$	[153]
		$2.2\,\text{fb}^{-1}$	CMS	$\sigma(tq)$, $\sigma(\bar{t}q)$, $\sigma(tq+\bar{t}q)$	[154]
		$2.3\,\text{fb}^{-1}$	CMS	$d\sigma/d\lvert y(t/\bar{t})\rvert$	[155]
		$35.9\,\text{fb}^{-1}$	CMS	$\sigma(tq)$, $\sigma(\bar{t}q)$, $\sigma(tq+\bar{t}q)$	[156]
Top decay	1.96 TeV	$2.7\,\text{fb}^{-1}$	CDF	F_0	[157]
		$8.7\,\text{fb}^{-1}$	CDF	F_0	[158]
		$5.4\,\text{fb}^{-1}$	D0	F_0	[159]
Top decay	7 TeV	$1.04\,\text{fb}^{-1}$	ATLAS	F_0, F_L	[160]
		$5.0\,\text{fb}^{-1}$	CMS	F_0, F_L	[161]
Top decay	8 TeV	$20.2\,\text{fb}^{-1}$	ATLAS	Γ_t	[162]
		$20.2\,\text{fb}^{-1}$	ATLAS	F_0, F_L	[163]
		$19.7\,\text{fb}^{-1}$	CMS	F_0, F_L	[164]
Top decay	13 TeV	$19.8\,\text{fb}^{-1}$	CMS	F_0, F_L	[165]

of the Wtb vertex in detail. In the search for BSM physics, single top-quark production processes are of particular interest, as they are expected to be sensitive to various potential BSM scenarios [168].

The single production of top quarks happens via different mechanisms, with the specific proportions of the individual processes depending on the energy scale. The t-channel is the dominant production channel at both the Tevatron and LHC, contributing approximately $70-80\,\%$ of the total single top-quark production cross section. In t-channel production, a virtual W boson and a b quark fuse to produce the top quark via $W^*b \to t$.

As the t-channel is the dominant production mechanism at current energies and has been measured very precisely, it is often subject to BSM and EFT studies [17, 23, 136, 169–171]. The t-channel production can be described in the *five-flavor scheme* (5FS) as a $2 \to 2$ process, or in the *four-flavor scheme* (4FS) as a $2 \to 3$ process. In the first case, the proton is assumed to contain a b sea quark, while in the latter case, also the b quark production via gluon splitting is included. It has been shown that calculations using both flavor schemes are in good agreement with each other [172]. The leading Feynman diagrams for t-channel production in both schemes are shown in Fig. 6.1.

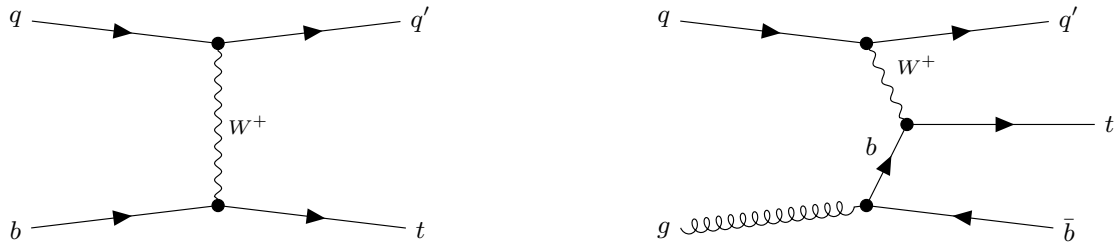

Figure 6.1: Leading order Feynman diagrams for t-channel single top-quark production in the 5FS (left) and in the 4FS (right).

Measurements of t-channel single top-quark production cross sections have been performed at Tevatron and LHC. In this study, measurements from the ATLAS and CMS experiments at 7, 8, and 13 TeV are considered, all unfolded to parton level. Specifically, the total single top-quark and single anti-top quark t-channel production cross sections, $\sigma(tq)$ and $\sigma(\bar{t}q)$, are included as well as differential cross sections as a function of the transverse momentum of the top quark, $\frac{\mathrm{d}\sigma}{\mathrm{d}p_T(t)}$, and of the anti-top quark $\frac{\mathrm{d}\sigma}{\mathrm{d}p_T(\bar{t})}$. Furthermore, differential cross section measurements as a function of the rapidity y of the (anti-) top quark are considered. If differential cross sections are published in terms of normalized distributions, the absolute distributions are reconstructed for this analysis using the associated total cross section measurements. The specific single top-quark production observables and measurements included in the analysis are listed in Tab. 6.1.

6.3.2 Top-quark decay observables

As mentioned in Sec. 4.3.3, due to its high mass, the top quark has a very short lifetime and the unique property of decaying before hadronization. The decay width Γ of the top quark is the inverse of its lifetime, $\Gamma = \tau^{-1}$. It is an interesting quantity to study the Wtb vertex structure and to search for BSM physics, as deviations from the SM predictions in this observable would hint at non-SM decay channels or modified top-quark couplings. The top-quark decay width has been measured with indirect methods by the D0 and CMS experiments [173, 174]. In the following, only the first direct measurement of the top-quark decay width by ATLAS, using $t\bar{t}$ events at 8 TeV, is included [162].[1]

As a consequence of its short lifetime, the spin information of the top quark does not decorrelate and is preserved in its decay products. Since the top quark decays almost exclusively via $t \to Wb$ into a b quark and a W boson, the polarizations of the W bosons are interesting observables and allow searching for non-SM effects. When choosing the rest frame of the top-quark, the W boson and the b quark obtain the same amount of momentum but with opposite directions. As the spin must be conserved, the W boson can take three different polarizations: left-handed, right-handed, or longitudinally polarized. The fractions $F_i = \Gamma_i/\Gamma$ of the respective polarizations are called W boson helicity fractions. A *leading order* (LO)

[1] A direct measurement of the top-quark decay width using the full Run-II data set has been published by ATLAS in Ref. [175] but is not included here.

QCD SM calculation using the assumption of the b quark being massless yields the following predictions for the W boson helicity fractions [176]:

$$F_0 = \frac{2m_t^2}{m_t^2 + 2m_W^2} \approx 0.7\,, \quad F_L = \frac{2m_W^2}{m_t^2 + 2m_W^2} \approx 0.3\,, \quad F_R = 0\,. \tag{6.4}$$

More precise calculations in *next-to-next-to-leading order* (NNLO) QCD [177] yield

$$F_0 = 0.687 \pm 0.005\,, \quad F_L = 0.311 \pm 0.005\,, \quad F_R = 0.0017 \pm 0.0001\,. \tag{6.5}$$

The fraction F_R being (almost) zero is a consequence of the W^+-boson and b quark both being right-handed in this case, which is highly suppressed in the SM due to the $V - A$ structure of the weak interaction. Measurements of the W boson helicity fractions have been performed by the ATLAS, CMS, D0, and CDF experiments. Here, measurements of F_0 at a center-of-mass energy of 1.96 TeV by CDF and D0, as well as measurements of F_0 and F_L at 7, 8, and 13 TeV by the ATLAS and CMS experiments are included. Again, the specific measurements considered in the analysis are given in Tab. 6.1.

6.4 Parametrizations of BSM contributions

As pointed out in Sec. 5.3, when using EFTfitter.jl for SMEFT fits, the dependences of the observables on the Wilson coefficients need to be formulated in terms of functions returning predictions for the values of the observables. In the following paragraphs, it is described how these dependences are determined for the observables considered here. In particular, it is demonstrated how computationally fast parametrizations are obtained that can be used for the Bayesian fits with EFTfitter.jl.

6.4.1 BSM contributions to single-top quark production

To determine BSM contributions to single-top quark t-channel cross sections as a function of dimension-six SMEFT operators, MC events are generated using MadGraph5_aMC@NLO [178] with the `dim6top_LO` [15] *Universal FeynRules Output* (UFO) model [179]. This allows computing the values of parton-level cross sections for arbitrary configurations of the Wilson coefficients in LO QCD. In the MC simulations, the BSM scale is set to $\Lambda = 1$ TeV, and only one insertion of a dimension-six operator is considered at a time. The top-quark mass is set to $m_t = 172.5$ GeV, and the renormalization and factorizations scales are set to $\mu_R = \mu_F = m_t$. The `MSTW2008lo` *parton distribution function* (PDF) set [180] is used. For the SM values of the total cross sections, *next-to-leading order* (NLO) QCD predictions are obtained using MadGraph5_aMC@NLO. Scale variation uncertainties are estimated by varying both the renormalization and factorization scales independently to $\mu = m_t/2$ and $\mu = 2m_t$. PDF uncertainties are estimated by varying the PDF sets, using the `CT10nlo` and `NNPDF23_nlo` PDF sets [181, 182]. The total theory uncertainties of the SM values are obtained by adding the statistical (MC), scale variation, and PDF uncertainties in quadrature. For differential cross sections, NLO QCD corrections are taken into account by applying k-factors using the NLO predictions in Refs. [148, 150, 152, 155].

In order to obtain continuous and computationally fast parametrizations for the dependences of the cross sections on the Wilson coefficients, interpolations of the MC cross sections are performed. The interpolations are based on the parametrization of cross sections as a function of the Wilson coefficients in Eq. (4.8), where the σ_i and σ_{ij} are the unknown parameters to be determined. In this analysis, two approaches for the interpolation are considered simultaneously. In the *linear ansatz*, only leading BSM contributions of the order $\mathcal{O}(\Lambda^{-2})$, originating from the interference of SM and BSM diagrams, are considered,

$$\sigma_{\text{MC}} = \sigma_{\text{SM}} + \sum_i \tilde{C}_i \sigma_i \quad \text{(linear)}\,. \tag{6.6}$$

As indicated in Sec. 4.3.3, it can be beneficial to also consider contributions $\mathcal{O}(\Lambda^{-4})$ that arise when squaring the BSM diagrams. When including these terms, the interpolation function has a quadratic dependence on the Wilson coefficients, and the *quadratic ansatz* reads

$$\sigma_{\text{MC}} = \sigma_{\text{SM}} + \sum_i \tilde{C}_i \sigma_i + \sum_{i \leq j} \tilde{C}_i \tilde{C}_j \sigma_{ij} \quad \text{(quadratic)}\,. \tag{6.7}$$

It should be noted again that in this thesis, terms of the order $\mathcal{O}(\Lambda^{-4})$ are considered only when coming from dimension-six operators. Possible contributions from dimension-eight operators at the same order are not taken into account. By comparing both interpolation approaches during the following analyses, the influence of quadratic terms on the results of the SMEFT fits is examined. In the case of differential distributions, the interpolations are performed separately for each bin.

An example of the resulting parametrizations of cross sections is depicted in Fig. 6.2, where the 7 TeV total single top-quark t-channel cross section is given as a function of the coefficients $\tilde{C}_{\varphi q}^{(3)}$, $\tilde{C}_{uW}$, and $\tilde{C}_{qq}$. The plots show slices of the phase space where only one of the Wilson coefficients is varied at a time while the others are fixed to zero. Both the linear (light blue) and the quadratic interpolation (dark blue) are shown. It is visible that at the intersections with the corresponding ATLAS measurement (gray), the sensitivities of both parametrizations are very similar.

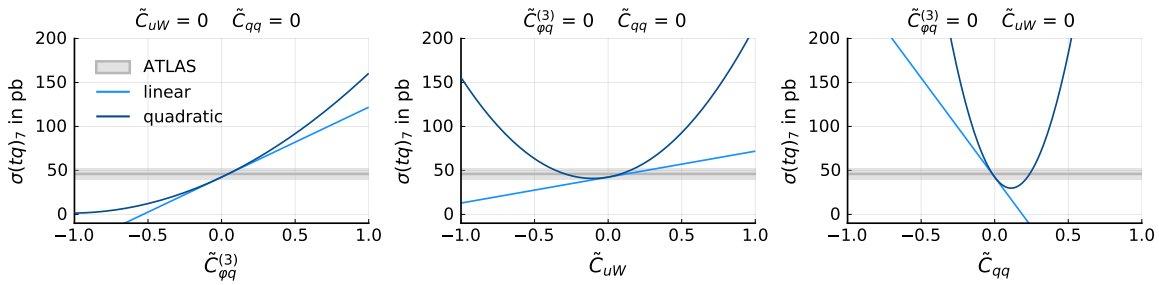

Figure 6.2: Dependence of the 7 TeV total single top-quark t-channel cross section on $\tilde{C}_{\varphi q}^{(3)}$, $\tilde{C}_{uW}$, and $\tilde{C}_{qq}$. Shown are slices of the phase space where only one of the Wilson coefficients is varied at a time while the others are fixed to zero. Both the linear (light blue) and the quadratic interpolation (dark blue) are shown. The ATLAS measurement of Ref. [148] and the corresponding total uncertainty are indicated in gray.

6.4.2 BSM contributions to top-quark decay observables

The dependence of the top-quark decay width and of the W boson helicity fractions on the Wilson coefficients is determined according to Ref. [183], considering only the operators O_{uW} and $O_{\varphi q}^{(3)}$, both modifying the Wtb vertex. BSM contributions are included at LO QCD, while for the SM value, NNLO QCD predictions are applied [177, 184]. In this analysis, the W boson helicity fractions are considered to depend only on $\tilde{C}_{uW}$, and the top-quark decay width is assumed to depend on $\tilde{C}_{uW}$ and $\tilde{C}_{\varphi q}^{(3)}$. In contrast to Ref. [183], for the W helicity fractions also quadratic contributions in $\tilde{C}_{uW}$ are taken into account. The dependence of the top-quark decay observables on the Wilson coefficients is visualized in Fig. 6.3.

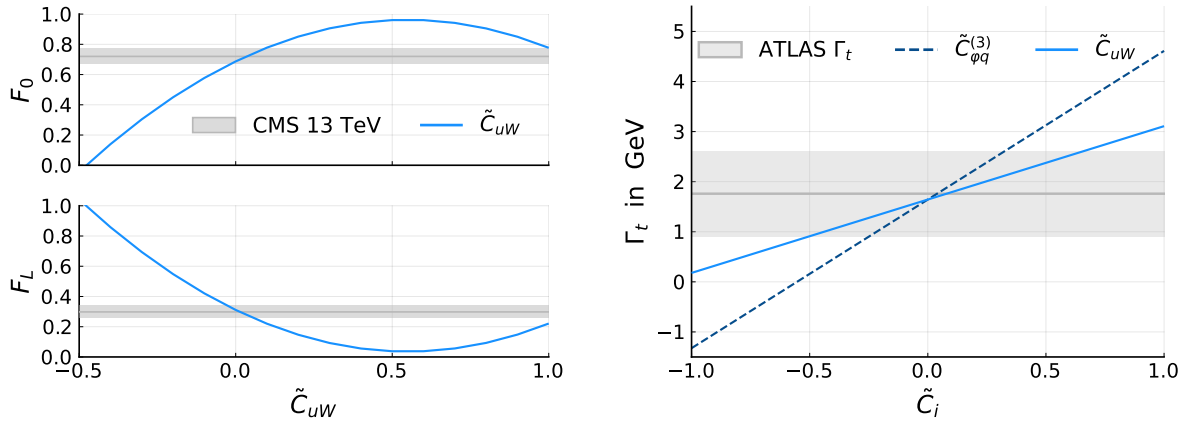

Figure 6.3: Dependence of the W boson helicity fractions F_0 and F_L, and of the top-quark decay width Γ_t on the Wilson coefficients $\tilde{C}_{uW}$ and $\tilde{C}_{\varphi q}^{(3)}$. For the decay width, only one Wilson coefficient is varied at a time while the other is fixed to zero. The corresponding measurements from ATLAS [148] and CMS [165], as well as the respective total uncertainties, are indicated in gray.

6.5 Uncertainties and correlation scenarios

Three different categories of uncertainties are considered in the fit: statistical uncertainties, systematic uncertainties, and theory uncertainties. For the systematic uncertainties, publications typically provide breakdowns of the different sources of these uncertainties. In this analysis, only the reported total systematic uncertainties are taken into account. The theory uncertainties are those of the SM predictions.

As detailed in Sec. 5.2, the total covariance matrix in EFTfitter.jl is constructed as the sum over the individual covariance matrices of the different types of uncertainties. Thus, correlation matrices have to be provided for each of the three uncertainty categories considered. As mentioned before, determining the correlations between multiple measurements from different experiments is an intricate task. As a consequence, the full correlation matrices for the measurements included here are not known. However, it has been pointed out, for example in Ref. [185], that measurements of single top quark cross sections from different experiments and at different energies can have non-negligible correlations.

Therefore, three different correlation scenarios are introduced to study their impact in the following:

- In the *no correlations* scenario, all measurements are considered to be uncorrelated, and each of the three correlation matrices is chosen to be the identity matrix.
- In the *known correlations* scenario, only the available correlation coefficients given in the publications are taken into account.
- In the *best guess correlations* scenario, based on simplifying assumptions, parametrizations of the correlation matrices are developed, and correlations between the measurements are introduced for studying their impact on the results of the fits.

In the following paragraphs, the correlation matrices of the different scenarios are discussed for the individual uncertainty categories.

Statistical uncertainties Correlations of statistical uncertainties are present when multiple measurements are based on the same data. As this is only the case for measurements performed at the same experiment and with the same data sets, correlations of the statistical uncertainties are usually estimated in the respective analyses and provided by the collaborations. Here, this is the case for the single top-quark t-channel cross section measurements, in particular for the correlations between the bins of differential distributions. Furthermore, correlation coefficients for the W boson helicity fractions F_0 and F_L are provided in the respective publications. Statistical correlations between measurements performed at different energies or at different experiments are not possible because independent events are used. The correlation coefficients given in the publications are used to construct the full 58×58 correlation matrix of the statistical uncertainties. The resulting matrix is used for both the *known correlations* and the *best guess correlations* scenarios, as no further statistical correlations are expected. The full correlation matrix of the statistical uncertainties is given in Fig. C.1 of App. C. The matrix contains only a very small number of non-zero elements and is almost diagonal, i.e., it differs only slightly from the identity matrix of the *no correlations* scenario.

Systematic uncertainties Correlations of systematic uncertainties arise between measurements that use the same experimental setup, the same data collection procedure, the same MC simulations, or the same data analysis methods. However, the estimation of systematic uncertainties is an elaborate and time-consuming task since usually, a variety of sources of systematic uncertainties has to be considered. Therefore, determining correlations of systematic uncertainties, especially between different experiments, is usually not feasible, and almost no information about such correlations is provided by the collaborations. In particular, it is not possible to estimate the correlations of systematic uncertainties outside the respective analysis teams. For these reasons, almost no estimates for systematic correlations between the considered measurements are available, with the exception of correlation coefficients for the measurements of F_0 and F_L, which are given in the respective publications. Therefore, the systematic correlation matrix of the *known correlations* scenarios is also very similar to the

identity matrix. In order to study the impact of potential systematic correlations on SMEFT fits, the following assumptions are made to construct a parametrization of the systematic correlation matrix to be used in the *best guess correlations* scenario:

- Measurements performed at the same experiment and at the same center-of-mass energy are assumed to be correlated with a correlation coefficient ρ_{sys} since for such measurements, the systematic uncertainties are likely to be coming from the same sources.

- Measurements performed at the same experiment but at different energies are assumed to be less correlated, as differences in the detectors, MC simulations, and energy-dependent systematic uncertainties are present. Such correlations are accounted for by a correlation coefficient $\rho_{\text{sys}}/2$.

- For measurements performed at different experiments, the systematic uncertainties are assumed to be uncorrelated.

Of course, these assumptions are oversimplifying, for example, because all bins of differential distributions are assumed to be correlated with the same correlation coefficient. However, the assumptions are largely based on the studies combining single top-quark cross section measurements of ATLAS and CMS in Ref. [185], and without further knowledge of the actual correlations, more specific assumptions cannot be justified. For the *best guess correlations* scenario, the correlation coefficient ρ_{sys} is chosen to be

$$\rho_{\text{sys}} = 0.9\,, \tag{6.8}$$

which is again based on the studies of Ref. [185]. The W boson helicity fractions F_0 and F_L are always assumed to be anti-correlated. The full 58×58 systematic correlation matrix based on the aforementioned assumptions is given in Fig. C.2. To illustrate the main aspects of the parametrization, a simplified representation is given in Eq. (6.9), considering only a subset of five different measurements. When taking into account only the total single top-quark and single anti-top quark cross sections measured by ATLAS at 7 TeV, $\sigma(tq)_7^{\text{A}}$ and $\sigma(\bar{t}q)_7^{\text{A}}$, the single top-quark cross sections at 8 TeV measured by ATLAS, $\sigma(tq)_8^{\text{A}}$, and CMS, $\sigma(tq)_8^{\text{C}}$, and the top-quark width measurement, Γ_t, the correlation matrix of the systematic uncertainties in the *best guess correlations* scenario is given as

$$M_{\text{sys}} = \begin{array}{c} \\ \sigma(tq)_7^{\text{A}} \\ \sigma(\bar{t}q)_7^{\text{A}} \\ \sigma(tq)_8^{\text{A}} \\ \sigma(tq)_8^{\text{C}} \\ \Gamma_t \end{array} \begin{array}{c} \begin{array}{ccccc} \sigma(tq)_7^{\text{A}} & \sigma(\bar{t}q)_7^{\text{A}} & \sigma(tq)_8^{\text{A}} & \sigma(tq)_8^{\text{C}} & \Gamma_t \end{array} \\ \begin{pmatrix} 1 & \rho_{\text{sys}} & \frac{\rho_{\text{sys}}}{2} & 0 & 0 \\ \rho_{\text{sys}} & 1 & \frac{\rho_{\text{sys}}}{2} & 0 & 0 \\ \frac{\rho_{\text{sys}}}{2} & \frac{\rho_{\text{sys}}}{2} & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}. \tag{6.9}$$

It should be noted that for $\rho_{\text{sys}} = 0$, the systematic correlations vanish and the *known correlations scenario* is obtained.

Theory uncertainties Also on the correlations between the theory uncertainties of the included measurements, no information is available. Therefore, simplifying assumptions are made in order to provide a *best guess correlations* matrix for these uncertainties. It is assumed that theory uncertainties are independent of the experiments and only depend on the center-of-mass energy:

- Measurements performed at the same center-of-mass energy are assumed to be correlated with ρ_{th}.
- Measurements of the same observable performed at different center-of-mass energies are assumed to be less correlated with $\rho_{\text{th}}/2$.

Again, these assumptions are rather simplistic, and the theory uncertainties usually depend on various factors, for example, on the energy scale of the collision. Such dependencies cannot be completely parametrized using only one parameter. Based on Ref. [185], for the *best guess correlations* scenario, the correlation coefficient ρ_{th} is chosen to be

$$\rho_{\text{th}} = 0.9\,. \tag{6.10}$$

The full correlation matrix of the *best guess correlations* scenario constructed using these assumptions is given in Fig. C.3. Similar to Eq. (6.9), the simplified representation of the systematic correlation matrix is given as:

$$M_{\text{th}} = \begin{array}{c} \\ \sigma(tq)_7^{\text{A}} \\ \sigma(\bar{t}q)_7^{\text{A}} \\ \sigma(tq)_8^{\text{A}} \\ \sigma(tq)_8^{\text{C}} \\ \Gamma_t \end{array} \begin{array}{c} \begin{array}{ccccc} \sigma(tq)_7^{\text{A}} & \sigma(\bar{t}q)_7^{\text{A}} & \sigma(tq)_8^{\text{A}} & \sigma(tq)_8^{\text{C}} & \Gamma_t \end{array} \\ \begin{pmatrix} 1 & \rho_{\text{th}} & \frac{\rho_{\text{th}}}{2} & \frac{\rho_{\text{th}}}{2} & 0 \\ \rho_{\text{th}} & 1 & \frac{\rho_{\text{th}}}{2} & \frac{\rho_{\text{th}}}{2} & 0 \\ \frac{\rho_{\text{th}}}{2} & \frac{\rho_{\text{th}}}{2} & 1 & \rho_{\text{th}} & 0 \\ \frac{\rho_{\text{th}}}{2} & \frac{\rho_{\text{th}}}{2} & \rho_{\text{th}} & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}. \tag{6.11}$$

6.6 Constraining the Wilson coefficients

The dependences of the observables on the Wilson coefficients determined in Sec. 6.4 are implemented into EFTfitter.jl, and the measurements of Tab. 6.1 are inserted together with their associated uncertainties. Uniform prior distributions are chosen for all three Wilson coefficients, with their maximum range being $-1 \leq \tilde{C}_i \leq 1$. For values $|\tilde{C}_i| > 1$, the EFT expansion of Eq. (4.6) breaks down, as the expansion parameter becomes to large and reaches perturbativity limits.

6.6.1 Constraints in the *known correlations* scenario

In a first step, the correlation matrices of the *known correlations* scenario are used in the fit. Since only a very small number of statistical correlation coefficients are available and

considered in this scenario, the results obtained are very similar to those of the *no correlation* scenario, which is therefore not discussed further here.

In Fig. 6.4, the resulting posterior distributions obtained from fits with EFTfitter.jl are shown for both the linear and the quadratic parametrizations of the single top-quark cross sections. Constraints on the three Wilson coefficients are shown in terms of one- and two-dimensional marginal distributions. On the main diagonal and in the upper triangle of the plots, the results of the total combination of all 58 measurements are shown, with the smallest intervals and areas containing 90 % of the posterior probability, i.e., the 90 % HDLs and HDAs, being highlighted. In the lower triangle, the 90 % HDAs are shown as obtained when only a certain subset of observables is used in the fit. In these plots, the constraints coming from top-quark decay observables are shown in gray. The constraints coming from the total single top-quark cross sections are shown in light blue, and the constraints coming from differential cross section measurements are shown in medium blue. The results of the total combination of all measurements are shown in dark blue and are the same as the ones in the upper triangle.

It is visible in the lower triangles of both pots that the top-quark decay measurements only constrain $\tilde{C}_{\varphi q}^{(3)}$ and $\tilde{C}_{uW}$. As already shown in Fig. 6.3, measurements of the W boson helicity fractions only constrain $\tilde{C}_{uW}$, while the top-quark decay width is also sensitive to $\tilde{C}_{\varphi q}^{(3)}$. When comparing the results based on the linear parametrization of the cross sections with those of the quadratic parametrization, it is observed that the shapes of the constraints obtained from the total and differential cross sections are different. In the case of the linear parametrizations, as expected from the likelihood in Eq. (5.9), a multivariate normal shape of the posterior distributions is observed. For the quadratic parametrization, the shapes of the HDAs are more distorted due to the $\mathcal{O}(\Lambda^{-4})$ BSM contributions to the observables. It can also be noticed that the constraints on $\tilde{C}_{\varphi q}^{(3)}$ and $\tilde{C}_{uW}$ coming from the total cross section measurements are much tighter in the case of the quadratic parametrization than in the linear one, which is also expected considering the sensitivity of the different parametrizations shown in Fig. 6.2. Overall, the total t-channel cross sections yield weaker constraints compared to those coming from the differential measurements, especially for $\tilde{C}_{\varphi q}^{(3)}$. Although the shapes and sizes of the constraints obtained from the total and differential cross sections differ between the two parametrizations, this is no longer observed in the total combination of all measurements. The results of the combination, shown in the upper triangles of Fig. 6.4, agree very well for both parametrizations. In the total combination, the coefficient $\tilde{C}_{qq}$ obtains the strongest constraints, with the width of its smallest 90 % interval being 0.026. The coefficient $\tilde{C}_{uW}$ is constrained to a similar degree and has a 90 % HDL of 0.029. The constraints on $\tilde{C}_{\varphi q}^{(3)}$ are the weakest, with the length of the 90 % interval being 0.074. All numerical values for the constraints are given in Tab. 6.2. There, the bounds of the marginal 90 % HDRs, their total widths (i.e., the 90 % HDLs), the modes, and the standard deviations (σ), as well as the deviations of the modes from the SM value in terms of the standard deviations are given for all three Wilson coefficients, both for the *known correlations* and the *best guess correlations* scenario. The values are determined using the linear parametrization of cross sections but are virtually the same as for the quadratic parametrization. The good agreement of the two different parametrizations of the cross sections in the total combination is also observed in

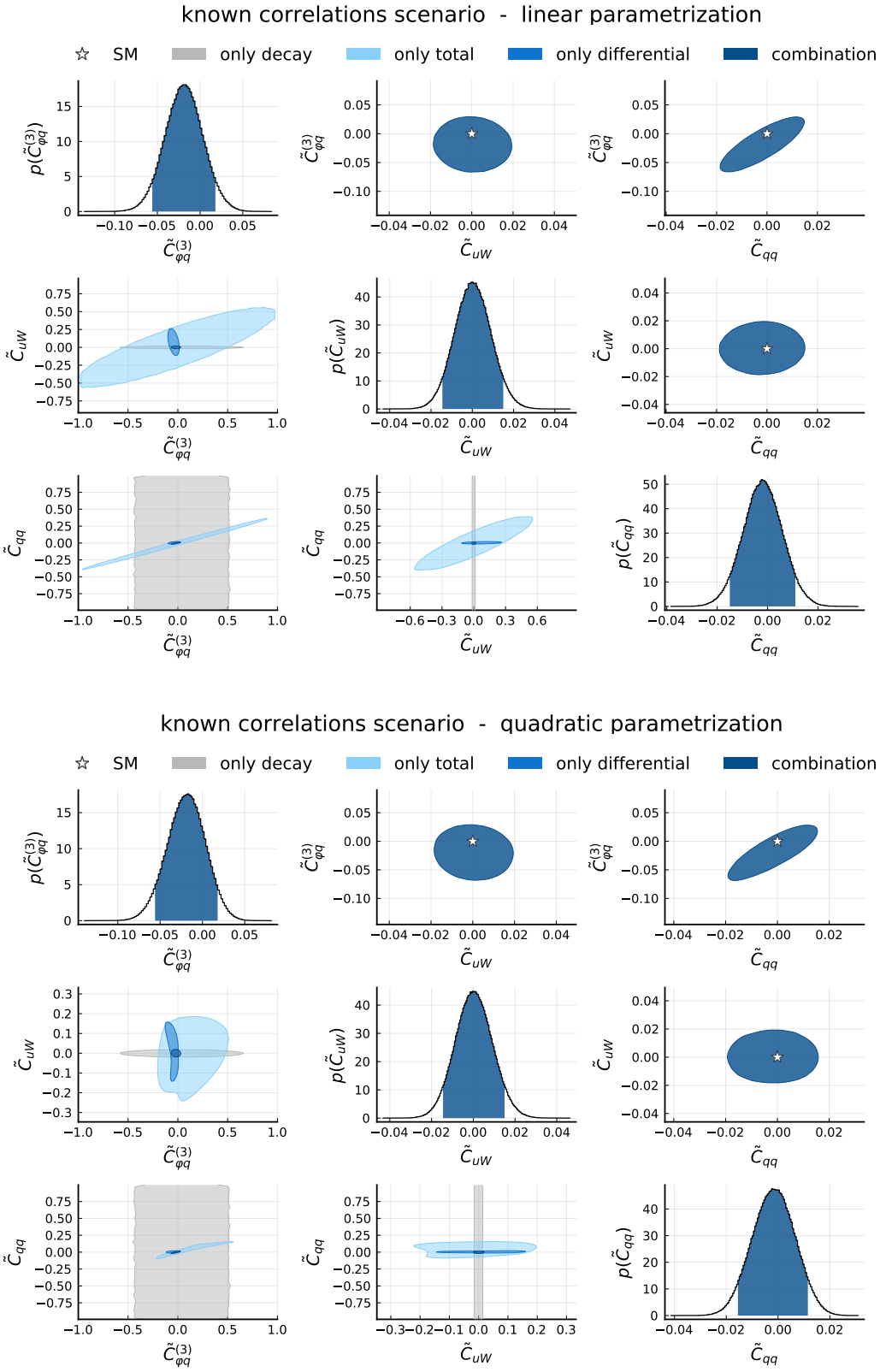

Figure 6.4: One- and two-dimensional marginalized posterior regions containing 90 % posterior probability obtained in fits using the *known correlations* scenario. In the lower triangle, the corresponding areas are shown for fits to the individual data sets. On the diagonal and in the upper triangle, the results of the total combination are shown.

the left plot of Fig. 6.7, where the constraints on the Wilson coefficients are visualized as the smallest 90 % intervals. In particular, it should be noticed that the SM value is contained within all the credible regions shown here.

To investigate the impact of the individual measurements on the combination in further detail, the MEWs, as introduced in Sec. 5.4, are computed. In Fig. 6.5, the resulting weights are shown for both the linear parametrization (in the top plot) and the quadratic parametrization (in the bottom plot). Following the discussions in Sec. 5.4, the question of whether the MEWs based on the sum of all two-dimensional HDAs are compatible with those based on the HDVs of the full three-dimensional parameter space is also examined. For this reason, in Fig. 6.5, the weights for both approaches are shown. The weights obtained from the three-dimensional 90 % HDVs are shown in light blue, while the weights based on the sum of all two-dimensional 90 % HDAs are shown in dark blue. It is noticeable in Fig. 6.5 that the general distribution of the weights and, in particular, their respective hierarchies are very similar for the linear and the quadratic parametrizations, with only a few exceptions.

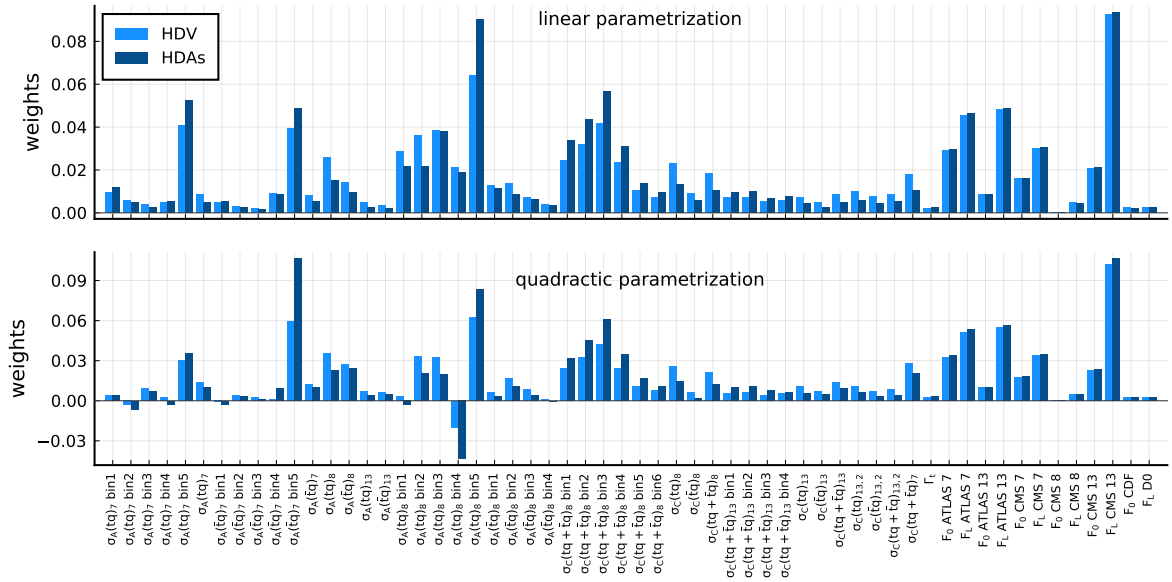

Figure 6.5: MEWs based on the 90 % HDV (light blue) and the sum of the 90 % HDAs (dark blue) for the linear (top) and the quadratic parametrization (bottom) of cross sections in the *known correlations* scenario.

While for the linear parametrization all measurements have positive weights, in the case of the quadratic parametrization, also negative weights are obtained, which indicates that excluding these measurements from the fit leads to a smaller 90 % HDR. This is particularly striking for the fourth bin of the ATLAS differential single top-quark cross section measurement at 8 TeV. It is assigned a rather large negative weight in the quadratic parametrization, while it obtains a positive weight in the linear parametrization. Comparing the weights obtained using the three-dimensional HDVs with those obtained using the sum over the two-dimensional HDAs, it can be observed that the overall distribution and hierarchy of weights are very similar in this case, again with a few exceptions. As the measurements are only weakly correlated

and the posterior distribution of the total combination has a very Gaussian shape, even for the quadratic parametrization, the similarity of the HDV and HDA weights is not surprising and agrees with the observations in the artificial example of Sec. 5.4, e.g., in Fig. 5.5. The largest weight is assigned to the measurement of the W helicity fraction F_L by CMS at 13 TeV. Most of the other measurements of F_L also receive large weights. This is mainly due to the tight constraints F_L yields on $\tilde{C}_{uW}$, which is also observed in Fig. 6.4. Other measurements with large weights are mostly those of differential cross sections, which is also plausible when considering the plots in the lower triangles of Fig. 6.4. The measurement of the top-quark decay width receives only a very small weight, as the constraints on $\tilde{C}_{\varphi q}^{(3)}$ provided by this measurement are much weaker than those from the total and differential cross section measurements. However, it should be noted again that the MEWs shown here are based only on measures of uncertainty, not on other influences of the measurements, such as their individual impacts on the location of the posterior distribution in the parameter space.

6.6.2 Constraints in the *best guess correlations* scenario

As the objective of this study is to investigate the impact of correlations on the constraints of Wilson coefficients, the Bayesian fit with EFTfitter.jl is repeated using the correlation matrices of the *best guess correlations* scenario. The resulting posterior distribution for the linear parametrization of cross sections is shown in Fig. 6.6. The one- and two-dimensional marginal distributions for the total combination are shown, as well as the 90 % HDAs of fits to the individual types of observables. The corresponding plot for the quadratic interpolation is given in Fig. C.4 of App. C.

Comparing the distributions in the lower triangle of Fig. 6.6 with those of the *known correlations* scenario in Fig. 6.4, it is visible that the main effect of the correlations in the *best guess correlations* scenario is a shift of the constraints coming from the differential cross sections. This shift is particularly large for $\tilde{C}_{\varphi q}^{(3)}$. It is noticeable in the lower triangle plots of Fig. 6.6 that due to the shift in $\tilde{C}_{\varphi q}^{(3)}$, the 90 % HDAs obtained using only the differential cross sections no longer have a common intersection with the constraints coming from the total cross sections and those coming from the W boson helicity fractions. Consequently, this also leads to a shift in the constraints of the total combination of all measurements. In particular for $\tilde{C}_{\varphi q}^{(3)}$, deviations from the SM expectation are obtained in the *best guess correlations* scenario, which is clearly visible in the distributions of the upper triangle, where the SM value lies outside the 90 % HDA for the $\tilde{C}_{\varphi q}^{(3)}$ axis. For the quadratic parametrization of cross sections, similar effects can be observed in Fig. C.4.

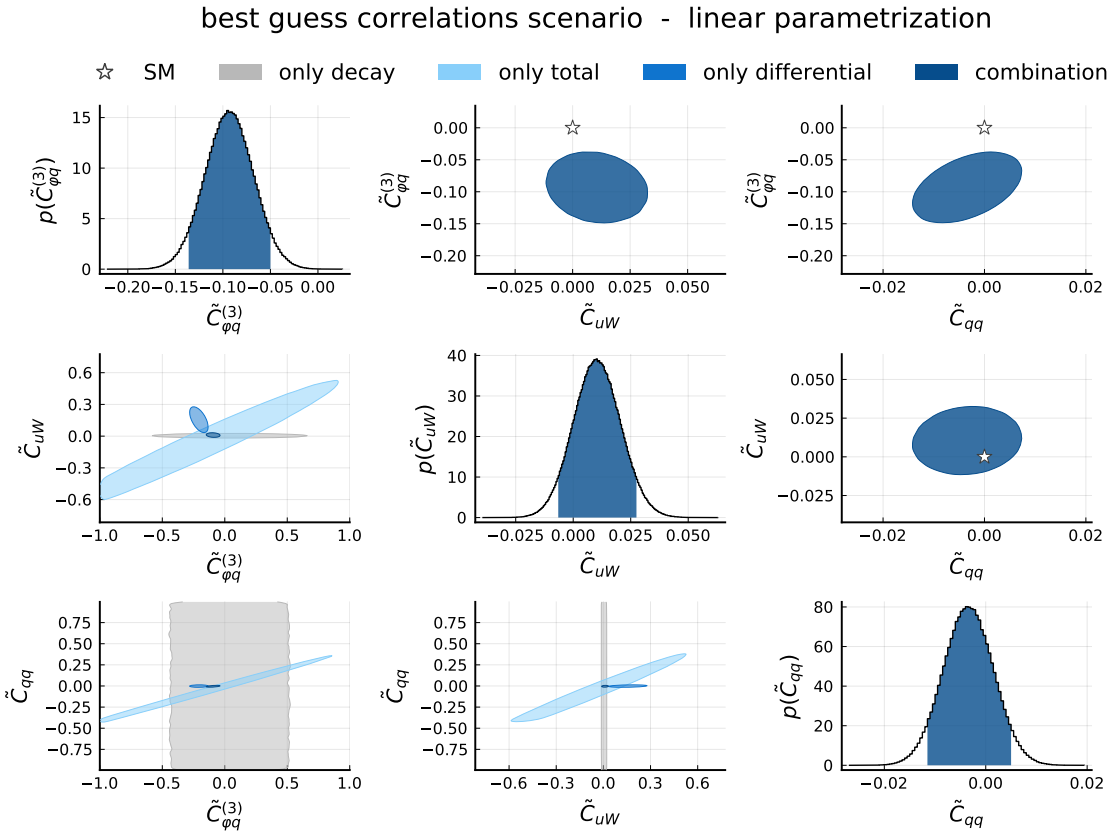

Figure 6.6: Posterior distribution of the *best guess correlations* scenario when using the linear parametrization of cross sections. Shown are the one- and two-dimensional marginalized posterior regions containing 90 % posterior probability. On the diagonal and in the upper triangle, the results of the total combination of all measurements are shown. In the lower triangle, the 90 % HDAs are shown for fits to the individual types of observables.

A clear visualization of the differences of the constraints obtained in the two correlation scenarios is given in Fig. 6.7, where the marginalized 90 % *highest density intervals* (HDIs) are shown for the *known correlations* scenario (left) and the *best guess correlations* scenario (right). The constraints are given for both the linear and the quadratic parametrizations of cross sections. The good agreement between the two parametrizations in the total combination is again observed. Also the shift of the constraints on $\tilde{C}_{\varphi q}^{(3)}$ away from the SM value towards negative values is clearly visible, with its mode shifting from -0.019 to -0.095, leading to a deviation from the SM value of about $3.65\,\sigma$ in the *best guess correlations* scenario. For $\tilde{C}_{uW}$, a minor shift towards larger values is observed in the *best guess correlations* scenario, with its mode changing from 0.00075 to 0.01025 and the 90 % HDR still being in agreement with the SM. For the coefficient $\tilde{C}_{qq}$, the mode shift is very minor. A decrease in the size of the 90 % HDI by about a factor of 1.6 is observed for the linear parametrization, while the corresponding intervals for $\tilde{C}_{\varphi q}^{(3)}$ and $\tilde{C}_{uW}$ increase by a factor of 1.15. This is also observed in the two-dimensional distributions of Fig. 6.6. However, such effects are expected, as the correlations lead to deformations of the multidimensional posterior distributions.

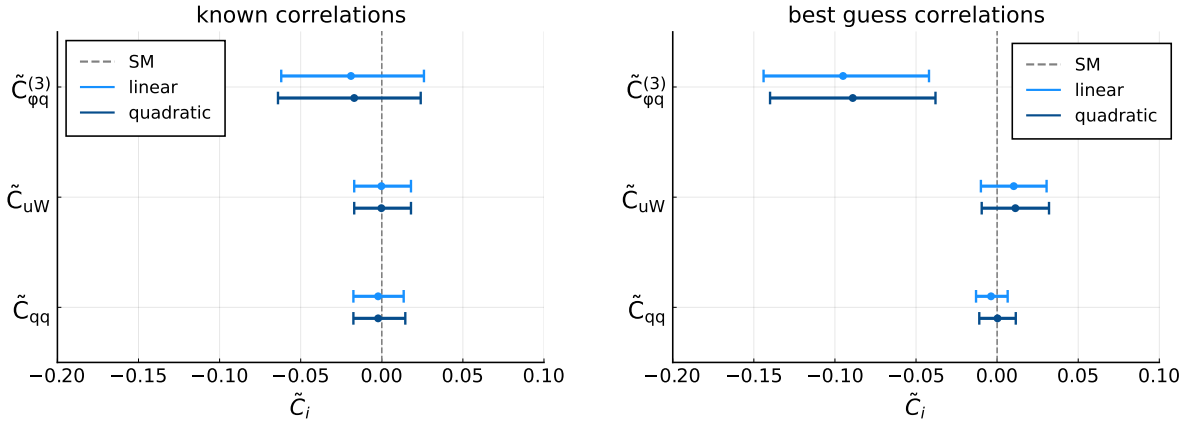

Figure 6.7: Marginal modes and smallest 90 % intervals of the one-dimensional marginalized posterior distributions using the linear (light blue) and quadratic (dark blue) parametrization of cross sections in the *known correlations* scenario (left) and in the *best guess correlations* scenario (right). The SM value is indicated as the dashed gray line.

The numerical values for the 90 % HDIs and the modes of the *best guess correlations* scenario are also given in Tab. 6.2.

Table 6.2: Bounds of the marginalized 90 % HDIs, their total length, the standard deviation, and the value of the mode, as well as the deviation of the mode from the SM value in units of the standard deviation.

	known correlations					*best guess correlations*				
$\tilde{C}_i$	90% HDI	90% HDL	σ	mode	mode/σ	90% HDI	90% HDL	σ	mode	mode/σ
$\tilde{C}^{(3)}_{\varphi q}$	[-0.056, 0.018]	0.074	0.023	−0.0190	−0.84	[-0.136, -0.051]	0.085	0.026	−0.0950	−3.65
$\tilde{C}_{uW}$	[-0.014, 0.015]	0.029	0.009	0.0008	0.08	[-0.007, 0.027]	0.034	0.011	0.0103	0.98
$\tilde{C}qq$	[-0.015, 0.011]	0.026	0.008	−0.0023	−0.29	[-0.012, 0.005]	0.016	0.005	−0.0038	−0.75

Also for the *best guess correlations* scenario, the MEWs are determined and shown in Fig. 6.8. In this scenario, the agreement between the hierarchies of the weights from the linear and the quadratic parametrizations is not as good as in the *known correlations* scenario. In general, for both parametrizations, more measurements are assigned negative weights than in the *known correlations* scenario. Again, in particular for the fourth bin of the ATLAS differential single top-quark cross section measurement at 8 TeV, a large negative weight is obtained in the quadratic parametrization, while a positive weight is obtained in the linear parametrization. However, also in this *best guess correlations* scenario, the weights based on the sum of the HDAs agree very well with those based on the HDVs.

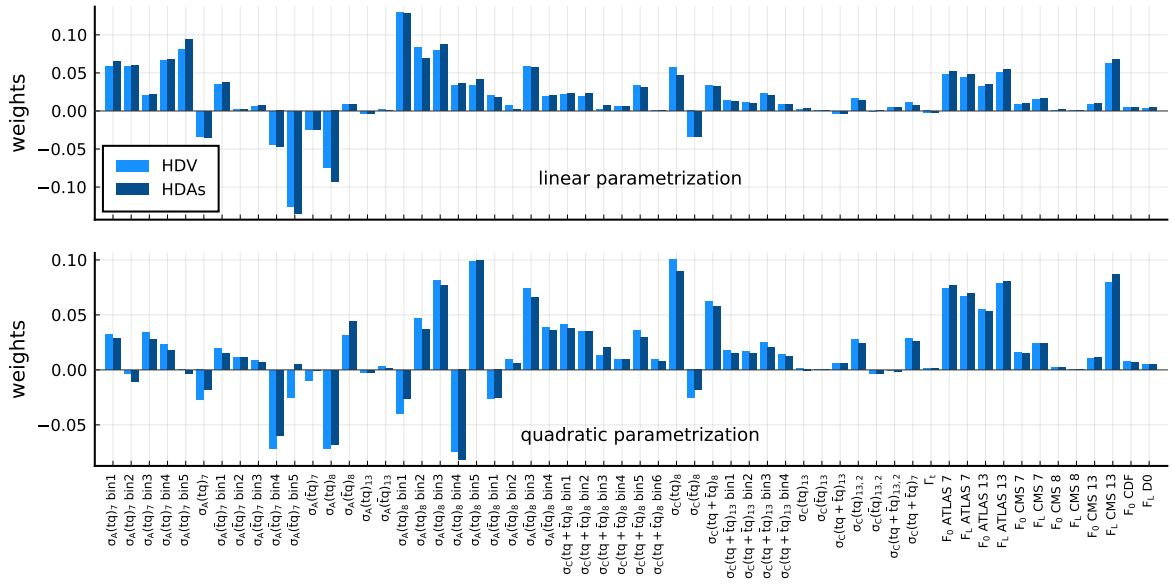

Figure 6.8: MEWs based on the 90 % HDV (light blue) and the sum of the 90 % HDAs (dark blue) for the linear (top) and the quadratic parametrization (bottom) of cross sections in the *best guess correlations* scenario.

As mentioned in Sec. 5.4, in EFTfitter.jl, the concept used to determine the MEWs can also be applied to compute weights for the different types of uncertainties, i.e., the UEWs, which are based on the relative decrease of the HDR size when excluding a particular type of uncertainty. To investigate the influence of the three types of uncertainties considered in this analysis, the UEWs are computed for both correlation scenarios. Again, both the three-dimensional HDV and the sum of the two-dimensional HDAs are used to determine these weights. In Fig. 6.9, the resulting weights are shown for the *known correlations* (left) and the *best guess correlations* (right) scenarios, both for the linear and the quadratic parametrizations.

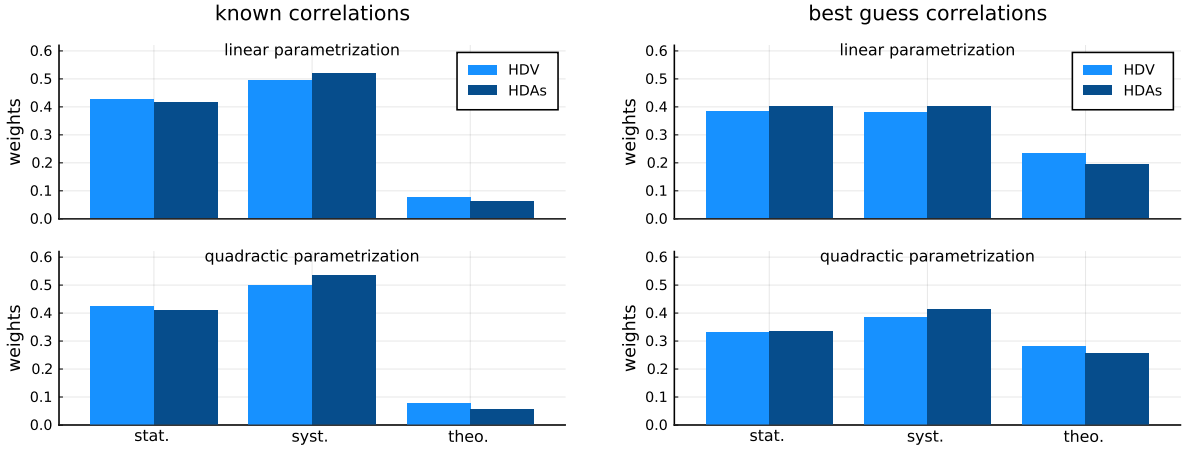

Figure 6.9: UEWs based on the 90 % HDV (light blue) and the sum of the 90 % HDAs (dark blue) for the linear (top) and the quadratic parametrization (bottom) in the *known correlations* scenario (left) and in the *best guess correlations* scenario (right).

Generally, a good agreement between the UEW for the linear and the quadratic parametrizations is observed. Also, the weights based on the HDV are very similar to those obtained from the sum of the HDAs, in particular considering the relative influence and hierarchies of the uncertainty types. For the fit in the *known correlations* scenario, the systematic uncertainties are the dominating type of uncertainty, with a relative weight of about 0.5. The statistical uncertainties are the second most important contributions, with a weight of about 0.4. The theory uncertainties are rather insignificant in this scenario and have a weight of less than 0.1. In the *best guess correlations* scenario, the statistical and the systematic uncertainties still obtain very similar weights between 0.3 and 0.4. The theory uncertainties, however, are more important in this scenario and obtain larger weights between 0.2 and 0.3. This can be explained by the large number of positive correlations assumed between the theory uncertainties in this *best guess correlations* scenario (see the correlation matrix in Fig. C.3). The impact of the uncertainties on the resulting constraints is further examined in Sec. 6.7.

6.6.3 Varying the correlation coefficients

The observed impact of correlations on the constraints is further investigated by considering the *best guess correlations* scenario in more detail. The parametrization of the correlation matrices introduced in Sec. 6.5 is still used but the values of the correlation coefficients ρ_{sys} and ρ_{th} are varied from those of the *best guess correlations* scenario, where they were chosen to be $\rho_{\text{sys}} = \rho_{\text{theo}} = 0.9$.

In a first step, only one of the correlation coefficients is varied at a time, or both are simultaneously set to the same value. In Fig. 6.10, the results obtained when varying the values of ρ_{sys} and ρ_{theo} between 0 and 0.95 in steps of 0.05 are shown for the Wilson coefficient $\tilde{C}_{\varphi q}^{(3)}$. Depicted are the marginal mode and the 90 % HDR when using the linear parametrization. They are shown for the cases in which only one of the correlation coefficients is varied at a time while the other one is fixed to zero and for the case in which both correlation coefficients are varied simultaneously to the same value.

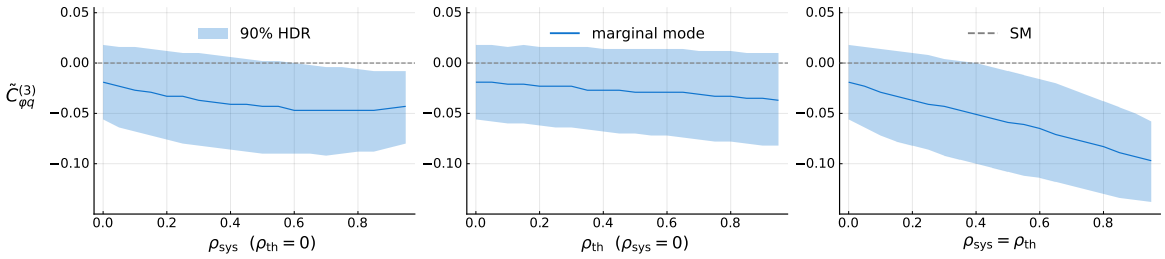

Figure 6.10: Marginal mode and smallest 90 % interval for $\tilde{C}_{\varphi q}^{(3)}$ when varying only ρ_{sys} (left), when varying only ρ_{th} (middle), and when varying both simultaneously (right) from 0 to 0.95 in steps of 0.05.

In all three cases, a continuous behavior of the changes of the mode and of the 90 % HDR is observed. The shifts of the mode and the interval towards more negative values are again visible, with the deviations from the SM increasing for larger values of the correlation coefficients.

It is noticeable that this effect is also present when only one of the correlation coefficients is varied but is much stronger when both correlation coefficients are varied simultaneously. The corresponding plots for the other two Wilson coefficients are shown in Fig. C.5 and Fig. C.6 of App. C. There, the impact of the correlations is not as strong as for $\tilde{C}_{\varphi q}^{(3)}$, but it is also continuous and in agreement with the observations in Fig. 6.7. A very similar behavior is observed when using the quadratic parametrization, and the corresponding plots are shown in App. C.3.2.

In a next step, the values of the correlation coefficients ρ_{sys} and ρ_{th} are varied separately. In Fig. 6.11, this is shown for $\rho_{\text{sys}}, \rho_{\text{theo}} \in [0, 0.15, 0.3, 0.45, 0.6, 0.75, 0.9]$. The relative deviation of the marginal mode towards negative values (left) and the relative increase of the 90 % HDL (right) compared to the *known correlations* scenario are shown for the Wilson coefficient $\tilde{C}_{\varphi q}^{(3)}$. In these plots, the top left bin corresponds to the *known correlations* scenario, and the bottom right bin represents the *best guess correlations* scenario.

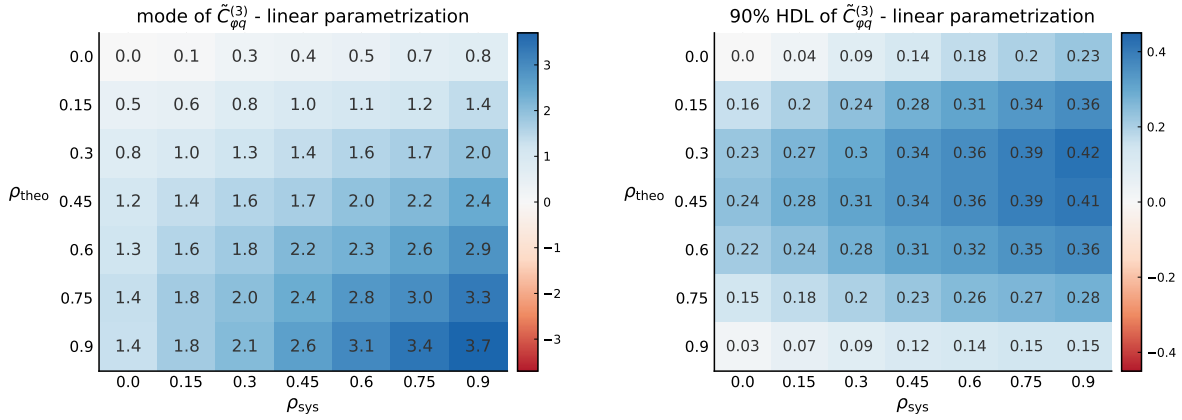

Figure 6.11: Relative deviation of the marginal mode (left) and relative increase of the 90 % HDL compared to the *known correlations* scenario for $\tilde{C}_{\varphi q}^{(3)}$ when varying the values of ρ_{sys} and ρ_{th} independently.

It is observed that variations of both correlation coefficients lead to deviations of the mode towards more negative values and that especially for large values of both correlation coefficients, these deviations from the *known correlations* scenario increase. For the 90 % HDL it is noticeable that it steadily increases for larger values of ρ_{sys}. When varying ρ_{th}, however, the HDL reaches its maximum for medium values of $\rho_{\text{th}} \approx 0.3$. The corresponding plots for $\tilde{C}_{uW}$ and $\tilde{C}_{qq}$ are shown in Fig. C.10 and Fig. C.11, respectively. There, a similar behavior of the mode is observed for $\tilde{C}_{uW}$, while for $\tilde{C}_{qq}$ the largest deviation of the mode is found for medium values of $\rho_{\text{th}} \approx 0.3$. While the 90 % HDL of $\tilde{C}_{uW}$ increases only slightly for larger values of the correlation coefficients, for $\tilde{C}_{qq}$ a decrease of the HDL is observed, which reaches its maximum in the *best guess correlations* scenario.

6.7 Impact of correlations in future scenarios

The correlation studies above are based on measurements performed at the Tevatron and the LHC, in the latter case using only Run-I and preliminary Run-II data with a maximum integrated luminosity of $20.2\,\mathrm{fb}^{-1}$, as given in Tab. 6.1. During Run-II of the LHC, data with an integrated luminosity of about $140\,\mathrm{fb}^{-1}$ have been collected by both the ATLAS and the CMS experiments [186], and the full data sets are still to be analyzed with respect to the observables of interest. Currently, the LHC and its experiments are being prepared for Run-III, after which the amount of recorded data is expected to double to an integrated luminosity of about $300\,\mathrm{fb}^{-1}$ [187]. After this run, the LHC and the experiments will be upgraded once again to form the *High-Luminosity LHC* (HL-LHC), which is foreseen to start taking data at the end of this decade and to increase the amount of collected data by one order of magnitude, with an expected integrated luminosity of up to $3000\,\mathrm{fb}^{-1}$ [187]. With that much more data becoming available in the next years, the statistical uncertainties of the measurements are expected to decrease significantly. Therefore, it is a particularly interesting question how the constraints on the Wilson coefficients will be affected by the decreasing statistical uncertainties and what the impact of correlations will be in such future scenarios.

To investigate this, constraints on the three Wilson coefficients are again derived for both correlation scenarios but assuming improved precision of the measurements. For this purpose, a simple approximation of the expected precision is performed by scaling the statistical uncertainties of all measurements in Tab. 6.1 to an integrated luminosity of $300\,\mathrm{fb}^{-1}$, as expected after LHC Run-III . The systematic and theory uncertainties, as well as the nominal values of the measurements, are assumed to remain unchanged and are set to the same values as in the previous studies. The constraints on the Wilson coefficients with such reduced statistical uncertainties are derived again using the described setup of EFTfitter.jl. The resulting one-dimensional marginalized intervals containing 90 % of the posterior probability are shown in Fig. 6.12 for both the linear and the quadratic parametrization using the *known correlations* scenario (left) and the *best guess correlations* scenario (right). For a simpler comparison, also the corresponding results using the current statistical uncertainties, as shown in Fig. 6.7, are indicated again using slightly transparent markers.

It is visible that, as expected, the sizes of the 90 % intervals decrease for all three Wilson coefficients when assuming the reduced uncertainties. In the *known correlations* scenario, this effect is largest for $\tilde{C}_{uW}$, where the 90 % HDL shrinks by about a factor of two. Also the modes of all three Wilson coefficients shift slightly towards negative values, with the corresponding 90 % intervals still containing the SM value. When considering the *best guess correlation* scenario, there are much larger shifts in the modes of the posterior distributions compared to the results with the current statistical uncertainties. This stronger impact of the correlations in this scenario with reduced statistical uncertainties is expected, as the rather weakly correlated statistical uncertainties have less influence, and the stronger correlated systematic and theory uncertainties dominate. In particular for $\tilde{C}_{\varphi q}^{(3)}$, the shift of the mode towards more negative values, as well as a shrinkage of the 90 % HDL to about 70 % of its original length, is striking. In this scenario, the mode of $\tilde{C}_{\varphi q}^{(3)}$ deviates by about $9.2\,\sigma$ from the SM value. Also for $\tilde{C}_{uW}$, the

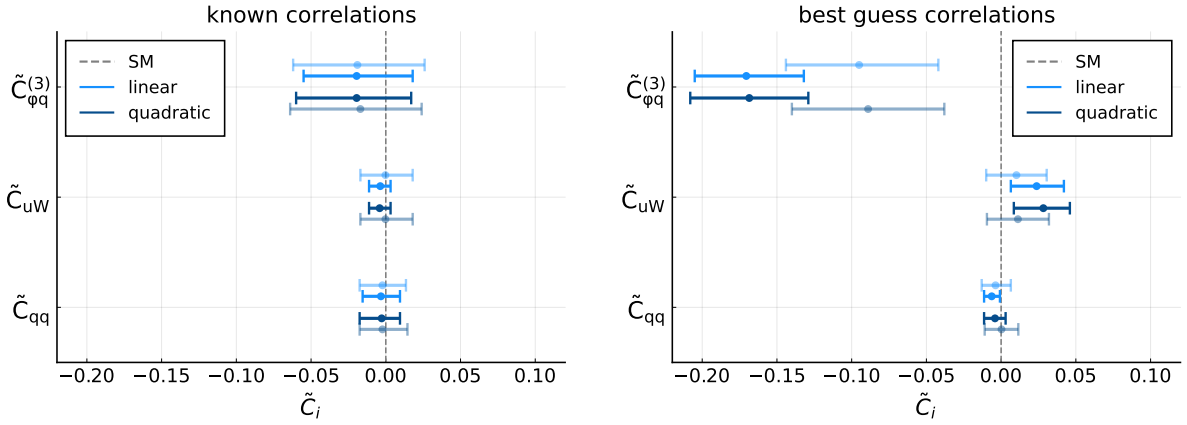

Figure 6.12: Marginal modes and smallest 90 % intervals of the one-dimensional marginalized posterior distributions using the linear (light blue) and quadratic (dark blue) parametrization of cross sections in the *known correlations* scenario (left) and in the *best guess correlations* scenario (right) for the statistical uncertainties scaled down to an integrated luminosity of 300 fb^{-1}. The SM value is indicated as the dashed gray line. The results using the current statistical uncertainties are indicated using the slightly transparent markers.

90 % credible interval is no longer in agreement with the SM when using the reduced statistical uncertainties. The mode is shifted towards larger positive values and deviates by about 2.6 σ from the SM value. The intervals of $\tilde{C}_{qq}$ shrink by a factor of 1.8, with the mode being slightly shifted towards negative values. While for the quadratic parametrization the 90 % interval of $\tilde{C}_{qq}$ is still in agreement with the SM, in the case of the linear parametrization, the interval does no longer include the SM value, with the mode deviating by about 2.3 σ. All numerical values for the bounds and widths of the 90 % HDIs, the modes, and standard deviations of this scenario with the statistical uncertainties scaled to 300 fb^{-1} are given in Tab. 6.3.

In Fig. 6.13, the marginal mode and the 90 % HDL of $\tilde{C}^{(3)}_{\varphi q}$ are shown as a function of the correlation coefficients ρ_{sys} and ρ_{th} when they are varied independently or set simultaneously to the same value. It can be noticed that in the case of reduced statistical uncertainties, the impact on the constraints when varying both correlation coefficients at a time is much larger than in the case of the current uncertainties shown in Fig. 6.10. The corresponding plots for $\tilde{C}_{uW}$ and $\tilde{C}_{qq}$ are given in Fig. C.12 and Fig. C.13.

Table 6.3: Bounds of the marginalized 90 % HDIs, their total length, the standard deviation, and the value of the mode, as well as the deviation of the mode from the SM value in units of the standard deviation for the measurements when assuming the statistical uncertainties reduced to 300 fb^{-1}.

	known correlations					*best guess correlations*				
$\tilde{C}_i$	90% HDI	90% HDL	σ	mode	mode/σ	90% HDI	90% HDL	σ	mode	mode/σ
$\tilde{C}^{(3)}_{\varphi q}$	[-0.050, 0.012]	0.061	0.019	−0.020	−1.04	[-0.200, -0.139]	0.061	0.019	−0.171	−9.22
$\tilde{C}_{uW}$	[-0.010, 0.002]	0.012	0.004	0.004	−1.01	[0.009, 0.040]	0.030	0.009	0.024	2.64
$\tilde{C}qq$	[-0.013, 0.007]	0.021	0.006	−0.003	−0.52	[-0.011, -0.002]	0.009	0.003	−0.006	−2.33

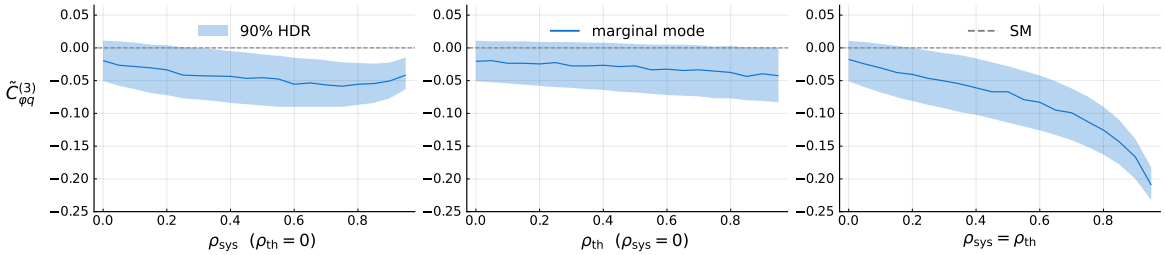

Figure 6.13: Marginal mode and smallest 90 % interval for $\tilde{C}^{(3)}_{\varphi q}$ when varying only ρ_{sys} (left), when varying only ρ_{th} (middle), and when varying both simultaneously (right) assuming the reduced statistical uncertainties.

Considering the UEWs for this future scenario, shown in Fig. 6.14, also confirms that the statistical uncertainties have much less influence on the results of the combination. In the *known correlations* scenario, the weight of the statistical uncertainties is now of a similar size as that of the theory uncertainties, both having weights of about 0.1. The influence of the systematic uncertainties is clearly dominating with a weight of about 0.8. Also in the *best guess correlations* scenario, the weight of the statistical uncertainties is around 0.1, and the fit is dominated by the systematic and theory uncertainties.

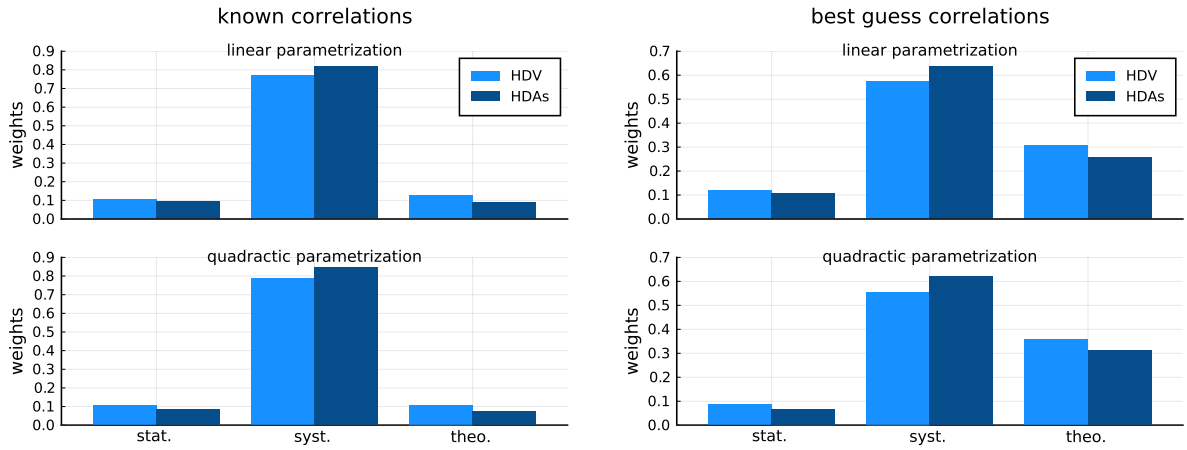

Figure 6.14: UEWs based on the 90 % HDV (light blue) and the sum of the 90 % HDAs (dark blue) for the linear (top) and the quadratic parametrization (bottom) in the *known correlations* scenario (left) and the *best guess correlations* scenario when assuming the reduced statistical uncertainties, corresponding to an integrated luminosity of 300 fb^{-1}.

When decreasing the statistical uncertainties further to a luminosity of 3000 fb^{-1}, which corresponds to the expected integrated luminosity of the HL-LHC, not much difference to the 300 fb^{-1} scenario is observed. This is expected, especially when considering the weights of the uncertainties in Fig. 6.14. The statistical uncertainties already have only a minor influence in this scenario, and the constraints are mainly driven by the systematic and theory uncertainties. However, it could be anticipated that during the next decades, until the HL-LHC stops operating, not only the statistical uncertainties on the measurements decrease due to more data collected, but that also the systematic and theory uncertainties decrease, for example,

due to improved detectors and more precise theory calculations and MC simulations. Here, a very simplistic estimation of such a future scenario is performed by scaling the statistical uncertainties of the measurements in Tab. 6.1 to an integrated luminosity of $3000\,\text{fb}^{-1}$ and simultaneously assuming an ambitious reduction of the systematic and theory uncertainties by a factor of two, for the purpose of estimating the effects of correlations on SMEFT fits in such a scenario. The resulting constraints and the corresponding numerical values are given in App. C in Fig. C.14 and Tab. C.1. Overall, a similar effect as for the $300\,\text{fb}^{-1}$ case can be observed. However, due to the reduced uncertainties, the obtained constraints are even tighter and therefore also lead to even larger deviations from the SM value. Again, in the *best guess correlations* scenario, BSM physics would be observed. The mode of $\tilde{C}_{\varphi q}^{(3)}$ deviates by more than $19\,\sigma$ in this case, and also for $\tilde{C}_{uW}$ and $\tilde{C}_{qq}$ deviations from the SM value by about $5\,\sigma$ are obtained. In the *known correlations* scenario, however, the constraints are still in agreement with the SM, with the deviations being only about $2\,\sigma$.

6.8 Summary

The studies presented in this chapter reveal that in SMEFT fits of several observables and multiple measurements, correlations between the uncertainties of measurements can have a significant impact on the results. It was demonstrated that the constraints on the Wilson coefficients depend on the correlation matrices and change for different values of the correlation coefficients. In particular, it was pointed out that for reduced uncertainties, which are expected from future measurements, the influence of correlations becomes even stronger and more important. It was pointed out that in such cases, the correlation matrices can be the crucial components of SMEFT fits that decide whether BSM phenomena are discovered or not. Generally, it should therefore be attempted to determine as many correlation coefficients as possible and to include them in the analyses, in particular when aiming towards global SMEFT fits using many different observables and measurements. In cases where this is not possible, the above studies provide guidance on how to investigate the influence of unknown correlation coefficients. By using knowledge about the underlying physics and measurements, a *best guess correlations* scenario can be constructed, similar to how it is done here. This scenario then allows studying the impact of potential correlations on the results of the particular fit. Different parametrizations of the correlation matrices and different values of the individual correlation coefficients can be tested. While in some cases, as in the studies above, a strong dependence of the results on the correlation matrices will be observed, this approach can, in particular, identify cases in which the correlations do not have a significant impact, which then allows validating the results. In any case, correlations between the uncertainties of measurements are essential components of each SMEFT fit that should be given particular attention.

7 Enhancing SMEFT fits by combining top-quark and B physics observables

The combination of different observables is a fundamental approach for improving the constraints on Wilson coefficients, as also demonstrated in the previous chapter. In this chapter, the opportunity of enhancing SMEFT fits by combining observables from top-quark and flavor physics is investigated. The feasibility of this approach is demonstrated, and the steps required for such a combination are outlined. For this purpose, two observables from top-quark and B physics are combined to constrain the Wilson coefficients of three SMEFT dipole operators modifying the top-quark coupling to the photon. This chapter is organized as follows: In Sec. 7.1, the different effective theories used to describe BSM contributions in top-quark and B physics are introduced, and the steps necessary for combining measurements from different energy scales in a common EFT approach are discussed. The observables and measurements included in this study are introduced in Sec. 7.2, and the computation of BSM contributions from dimension-six operators to the observables is detailed in Sec. 7.3. In Sec. 7.4, the constraints obtained in SMEFT fits to the individual and the combined observables are discussed, and the improvements gained through their combination are highlighted. In Sec. 7.5, the results of the study are briefly summarized.

7.1 Combining EFTs from different energy scales

As indicated before, SMEFT fits combining multiple observables and measurements from the top-quark sector gained much popularity over the last years and are a prominent approach for indirect BSM searches [3, 15–22]. While such combinations of different top-quark observables are already common, the combination with observables from other sectors of particle physics is particularly interesting, as further observables can provide additional sensitivity to the Wilson coefficients [170, 171]. Especially B physics processes offer great opportunities for improving the constraints on Wilson coefficients and may serve as a window to BSM physics, for example, when considering the observed anomalies discussed in Sec. 4.2 [188–192].

In this study, it is investigated how the Wilson coefficients of dimension-six SMEFT operators modifying top-quark couplings can be constrained using combined data from the top-quark and B physics sectors. For the purpose of exploring the required steps and to highlight the potential benefits of such a combination, only measurements of two different observables, namely the $t\bar{t}\gamma$ production cross section and the $\bar{B} \to X_s\gamma$ branching ratio, where X_s is any hadronic state containing a strange quark, are considered here. They are employed to constrain Wilson coefficients of the three dipole operators

$$\begin{gathered} O_{uB} = (\bar{q}_L \sigma^{\mu\nu} u_R)\, \tilde{\varphi} B_{\mu\nu}\,, \quad O_{uG} = \left(\bar{q}_L \sigma^{\mu\nu} T^A u_R\right) \tilde{\varphi} G^A_{\mu\nu}\,, \\ O_{uW} = \left(\bar{q}_L \sigma^{\mu\nu} \tau^I u_R\right) \tilde{\varphi} W^I_{\mu\nu}\,, \end{gathered} \tag{7.1}$$

as defined in Eq. (4.9). These are the SMEFT operators giving the leading BSM contributions of order $\mathcal{O}(\Lambda^{-2})$ to the two processes considered here. Contributions from dipole operators with right-handed b quarks contribute to top-quark decay and $b \to s$ transitions via one-loop diagrams but are suppressed by a factor m_b/m_t. Four-quark operators are neglected here since $t\bar{t}$ production at the LHC is dominated by the gluon-gluon channel, which makes up about 90 % of all top-quark pairs produced at 13 TeV [193]. BSM contributions from O_{uW} in top-quark decays are explicitly allowed. In this analysis, all Wilson coefficients are assumed to be real-valued, as only the real parts of the operators give contributions at order $\mathcal{O}(\Lambda^{-2})$.

When combining observables from top-quark and B physics, it has to be considered that the corresponding measurements are performed at different energy scales and that the computation of BSM contributions is typically performed using different effective theories. For top-quark physics, the measurements are performed at the top-quark scale $\mu_t \approx m_t$, and the corresponding SMEFT is defined at the energy scale Λ, which is assumed to be well above the scale of the SM. The considered B physics processes happen at much lower energies $\mu_b \approx m_b$, and an EFT describing these processes and enabling BSM calculations can be constructed at an energy scale below the electroweak scale, as briefly mentioned in Sec. 4.3.1. The steps required to combine both approaches in a common fit are described in the following paragraphs.

7.1.1 The weak effective theory

For BSM calculations of B physics processes, the WET is commonly used. In WET, the energy scale of the EFT expansion is at the scale $\mu_W \approx m_W$. The heavier particles, i.e., the heavy bosons, the top quark, and hypothetical heavy BSM particles, are integrated out. The Lagrangian of the WET is given as

$$\mathcal{L}_{\text{WET}} = \frac{4G_F}{\sqrt{2}} V_{ts}^* V_{tb} \sum_{i=1}^{10} \bar{C}_i Q_i \,, \tag{7.2}$$

where V_{ts} and V_{tb} are CKM matrix elements, and Q_i are the effective operators [125, 126]. The $\bar{C}_i$ are the corresponding Wilson coefficients of WET, which, in contrast to those of SMEFT, contain both the SM and BSM contributions. The operators relevant to describe the $b \to s$ transitions are:

$$\begin{aligned}
Q_1 &= (\bar{s}_L \gamma_\mu T^a c_L)(\bar{c}_L \gamma^\mu T^a b_L) \,, & Q_2 &= (\bar{s}_L \gamma_\mu c_L)(\bar{c}_L \gamma^\mu b_L) \,, \\
Q_3 &= (\bar{s}_L \gamma_\mu b_L) \sum_q (\bar{q} \gamma^\mu q) \,, & Q_4 &= (\bar{s}_L \gamma_\mu T^a b_L) \sum_q (\bar{q} \gamma^\mu T^a q) \,, \\
Q_5 &= (\bar{s}_L \gamma_\mu \gamma_\nu \gamma_\sigma b_L) \sum_q (\bar{q} \gamma^\mu \gamma^\nu \gamma^\sigma q) \,, & Q_6 &= (\bar{s}_L \gamma_\mu \gamma_\nu \gamma_\sigma T^a b_L) \sum_q (\bar{q} \gamma^\mu \gamma^\nu \gamma^\sigma T^a q) \,, \\
Q_7 &= \frac{e}{16\pi^2} m_b (\bar{s}_L \sigma^{\mu\nu} b_R) F_{\mu\nu} \,, & Q_8 &= \frac{g_s}{16\pi^2} m_b (\bar{s}_L \sigma^{\mu\nu} T^a b_R) G^a_{\mu\nu} \,, \\
Q_9^{ij} &= \frac{e^2}{16\pi^2} (\bar{s}_L \gamma_\mu b_L)(\bar{\ell}^i \gamma^\mu \ell^j) \,, & Q_{10}^{ij} &= \frac{e^2}{16\pi^2} (\bar{s}_L \gamma_\mu b_L)(\bar{\ell}^i \gamma^\mu \gamma_5 \ell^j) \,.
\end{aligned} \tag{7.3}$$

7.1.2 Matching SMEFT & WET

The WET provides a well-known framework for calculating SM and BSM contributions to B physics observables, and thus also for the $\bar{B} \to X_s\gamma$ branching ratio considered here. In this study, however, the B observables are used to derive constraints on the SMEFT coefficients $\tilde{C}_{uB}$, $\tilde{C}_{uG}$, and $\tilde{C}_{uW}$ at the top-quark scale. Therefore, the dependence of BSM contributions to $\mathrm{BR}(\bar{B} \to X_s\gamma)$ on the SMEFT Wilson coefficients needs to be determined. This is achieved by matching the SMEFT coefficients onto the WET coefficients [194]. For this, the different energy scales at which the two effective theories are defined have to be taken into account. In Fig. 7.1, a schematic representation of the energy scales of the different theories is shown. The steps for translating SMEFT coefficients into WET coefficients are indicated.

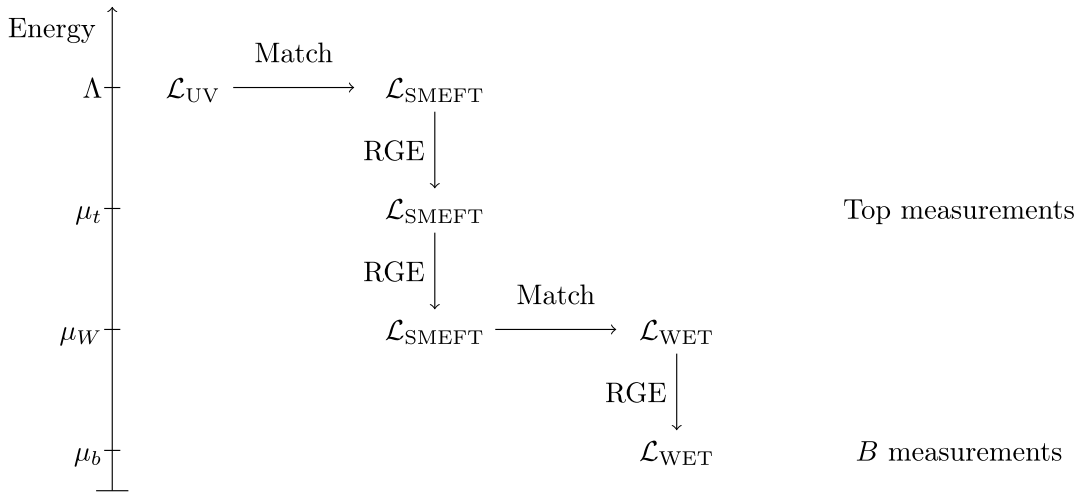

Figure 7.1: Schematic visualization of the energy scales of SMEFT and WET, and the steps necessary for translating SMEFT and WET coefficients [4].

At the energy scale of BSM physics, Λ, the underlying UV theory, which is the desired theory describing both SM and BSM physics, can be matched on the effective description of SMEFT. Using so-called *renormalization group equation* (RGE) evolution within the SMEFT framework, the Wilson coefficients at the lower energy scale $\mu_t \approx m_t$ can be derived. This is the scale where the top-quark measurements are located. In this thesis, all constraints on the SMEFT Wilson coefficients are given at this scale $\mu_t = 172.5\,\mathrm{GeV}$. When using B physics processes to constrain SMEFT Wilson coefficients at the scale μ_t, predictions for the BSM contributions to the B observables have to be determined in terms of the SMEFT coefficients at this energy scale. To translate the SMEFT coefficients into the WET coefficients at the lower energy scale $\mu_b \approx m_b$ where the B measurements are located, the following steps need to be applied, which then allow using the common WET framework to determine BSM contributions to the B observables:

1. RGE evolution in SMEFT is used to run down the energy scale from μ_t to the matching scale $\mu_W \approx m_W$. The RGEs for the SMEFT coefficients have been computed in Refs. [195–198]. In this thesis, the RGE evolution is performed at one-loop level using the `wilson` package [199].

2. At the scale μ_W, the SMEFT coefficients are matched onto the WET coefficients. The matching conditions for converting SMEFT into WET coefficients are derived in Ref. [191]. In this thesis, these conditions are applied, as detailed in Ref. [5]. The numerical values for the matching conditions employed here are given in App. D.1.

3. RGE evolution in WET is used to run down the coefficients to the scale of B physics, $\mu_b \approx m_b$. At this scale, the well-known WET formalism is applied to compute predictions for the values of the B observables, including BSM contributions. In this thesis, these steps are performed using the `wilson` and `flavio` [200] packages.

These three steps thus allow determining the dependence of B physics observables on the SMEFT Wilson coefficients at the scale μ_t, using the established WET framework for computing BSM contributions.

7.2 Observables & measurements

As mentioned above, only two observables are considered to be combined in this feasibility study. From the top-quark sector, it is the cross section of $t\bar{t}\gamma$ production, i.e., the associated production of a top-antitop quark pair with a photon. In Fig. 7.2, examples of leading-order Feynman diagrams of $t\bar{t}\gamma$ production are shown. In these diagrams, possible insertions of dimension-six operators modifying the vertices are indicated by the black dots.

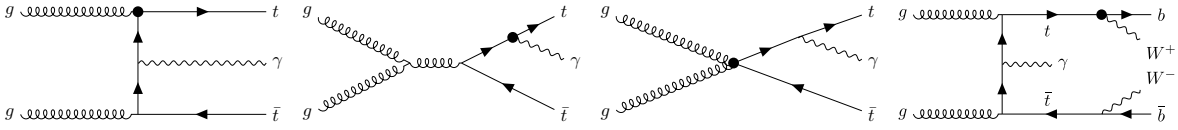

Figure 7.2: Examples of leading-order Feynman diagrams of $t\bar{t}\gamma$ production at the LHC. Possible insertions of dimension-six operators are indicated by the black dots.

Measurements of the $t\bar{t}\gamma$ production cross section have been performed at different center-of-mass energies by the ATLAS and CMS experiments [201–205]. Here, only the ATLAS measurement at 13 TeV, using an integrated luminosity of 36.1 fb^{-1}, of Ref. [201] is considered. There, the cross section is given as a fiducial cross section for final states containing one lepton (single-lepton channel) or two leptons (dilepton channel), where the leptons can be either electrons or muons (or the corresponding antifermions). The values of the measured cross sections, as well as the SM predictions in NLO QCD, are given in Tab. 7.1.

Table 7.1: Observables, measurements, and SM predictions considered in this study.

Observable		Measurement		SM prediction	Refs.
Top	$\sigma^{\text{fid}}(t\bar{t}\gamma \to 1\ell)$	ATLAS	$521 \pm 9\,(\text{stat.}) \pm 41\,(\text{syst.})$ fb	495 ± 99 fb	[201, 206]
Top	$\sigma^{\text{fid}}(t\bar{t}\gamma \to 2\ell)$	ATLAS	$69 \pm 3\,(\text{stat.}) \pm 4\,(\text{syst.})$ fb	63 ± 9 fb	[201, 206]
B	$\text{BR}(\bar{B} \to X_s\gamma)$	HFLAV	$(332 \pm 15) \times 10^{-6}$	$(336 \pm 23) \times 10^{-6}$	[207, 208]

From the B physics sector, measurements of $b \to s\gamma$ transitions are combined with the $t\bar{t}\gamma$ cross section. As visible in the Feynman diagrams in Fig. 7.3, such transitions are mainly possible via top-quark loops.[1] Therefore, similar vertices as in the top-quark processes of Fig. 7.2 are present, which makes these transitions sensitive to the same dimension-six operators. This, in particular, motivates the study and is the reason why the combination of these observables is expected to improve the constraints on the SMEFT Wilson coefficients. It should be noted that the sensitivity of the $b \to s\gamma$ transitions to $\tilde{C}_{uG}$ is induced by the SMEFT and WET running only. The specific $b \to s\gamma$ observable considered here is the $\bar{B} \to X_s\gamma$ branching fraction, $\mathrm{BR}(\bar{B} \to X_s\gamma)$. It has been measured by the BaBar [209–211], Belle [212–214], and CLEO [215] experiments. A combination of these measurements has been performed by the *Heavy Flavor Averaging Group* (HFLAV) [216] and accounts for small differences in the respective phase spaces. Here, the latest combined value [207] is used for the fit, as given in Tab. 7.1. Corresponding SM predictions of the branching ratio are considered at NNLO QCD [208] and are also given in Tab. 7.1.

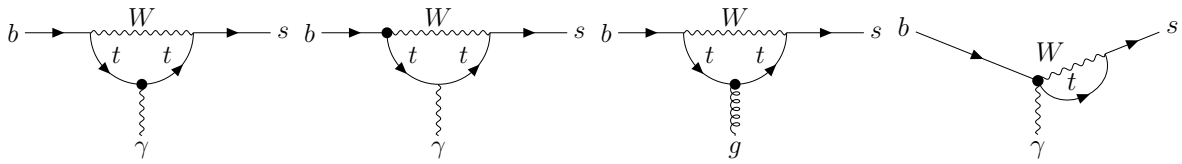

Figure 7.3: Examples of Feynman diagrams of $b \to s\gamma$ and $b \to sg$ transitions. Possible insertions of dimension-six operators are indicated as black dots.

7.3 Computing BSM contributions in SMEFT

7.3.1 $\mathrm{BR}(\bar{B} \to X_s\gamma)$

The dependence of $\mathrm{BR}(\bar{B} \to X_s\gamma)$ on the SMEFT Wilson coefficients $\tilde{C}_{uB}$, $\tilde{C}_{uG}$, and $\tilde{C}_{uW}$ is determined using the procedure described in Sec. 7.1. The BSM contributions are calculated at the scale μ_W in terms of WET coefficients using `flavio`, and the matching and running is performed using `wilson`. To obtain a fast parametrization of this dependence that can be used in EFTfitter.jl, a quadratic interpolation is performed. The resulting dependence of the $\bar{B} \to X_s\gamma$ branching ratio on $\tilde{C}_{uB}$, $\tilde{C}_{uG}$, and $\tilde{C}_{uW}$ is shown in Fig. 7.4, where one coefficient is varied at a time, while the others are set to zero.

It is visible that $\mathrm{BR}(\bar{B} \to X_s\gamma)$ is particularly sensitive to $\tilde{C}_{uB}$. Within the range considered, this coefficient agrees with the measurement in two regions, once at $\tilde{C}_{uB} \approx 0$ and once around $\tilde{C}_{uB} \approx -0.5$, due to the narrow parabolic shape of the dependence on this coefficient.

[1] In Fig. 7.3, also a diagram of a $b \to sg$ transition is shown, as it can be matched onto the WET coefficient $\bar{C}_8$, which due to running mixes with $\bar{C}_7$ and thus also contributes to the considered $b \to s\gamma$ transitions. More details on the running and matching of the WET and SMEFT operators can be found in Ref. [4].

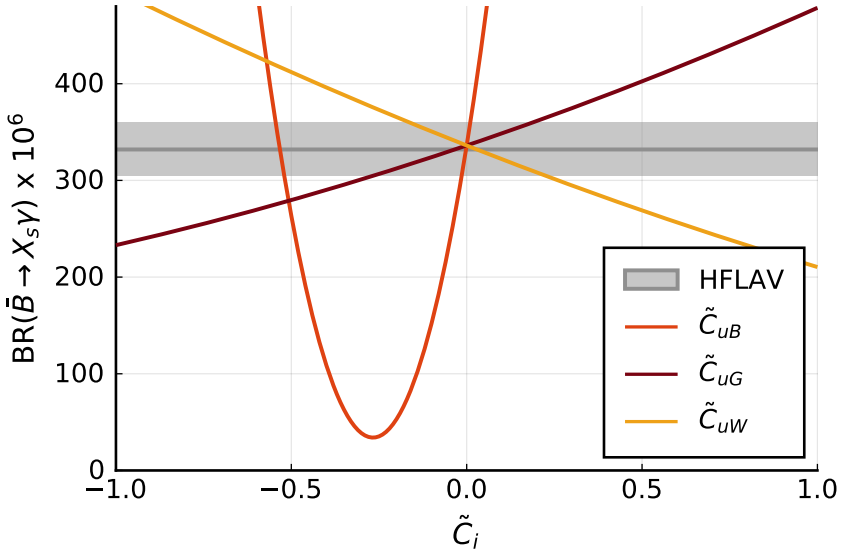

Figure 7.4: Dependence of BR($\bar{B} \to X_s\gamma$) on the SMEFT coefficients $\tilde{C}_{uB}$, $\tilde{C}_{uG}$, and $\tilde{C}_{uW}$ at the scale μ_t when only one of them is varied at a time, while the others are set to zero. The corresponding value of the HFLAV combination and the combined experimental and theoretical uncertainties are indicated by the gray line and band.

7.3.2 $t\bar{t}\gamma$ production

The parametrization of the $t\bar{t}\gamma$ cross section as a function of the SMEFT Wilson coefficients is determined using MC simulations with MadGraph5_aMC@NLO and the `dim6top_LO` UFO model, similar to Chap. 6. For $t\bar{t}\gamma$ production, however, only fiducial measurements are available, which requires additional steps in the computation, described in the following. The MC samples for $t\bar{t}\gamma$ are generated similarly to the signal samples of Ref. [201] as a $2 \to 7$ process. In the simulations, the BSM scale is chosen as $\Lambda = 1\,\text{TeV}$, and only one insertion of a dimension-six operator is allowed for each Feynman diagram. Since in the `dim6top_LO` UFO model different degrees of freedom are chosen for the SMEFT operators, it is not possible to directly set the value of $\tilde{C}_{uB}$, but only that of the linear combination

$$\tilde{C}_{uZ} = \cos\theta_W \tilde{C}_{uW} - \sin\theta_W \tilde{C}_{uB}\,, \tag{7.4}$$

with the Weinberg angle θ_W, where $\cos\theta_W = \frac{m_W}{m_Z}$. Therefore, the MC samples are generated in the space of $\tilde{C}_{uG}$, $\tilde{C}_{uW}$, and $\tilde{C}_{uZ}$, and the quadratic interpolation according to Eq. (4.8) is performed in this parameter space. Afterwards, in the EFTfitter.jl combination, the original parameter space of $\tilde{C}_{uB}$, $\tilde{C}_{uG}$, and $\tilde{C}_{uW}$ is used again by inverting Eq. (7.4). A total of 201 different sampling points in the three-dimensional space of Wilson coefficients is chosen, with up to two of the coefficients taking non-zero values at a time. At each sampling point, 50 000 events are generated in LO QCD. A parametrization of the total $t\bar{t}\gamma$ cross section as a function of $\tilde{C}_{uG}$, $\tilde{C}_{uW}$, and $\tilde{C}_{uZ}$ is determined using an interpolation according to Eq. (4.8). In Fig. 7.5, the sampling points and the result of the interpolation are shown as slices of the parameter space where only one of the coefficients is varied, while the others are fixed to zero.

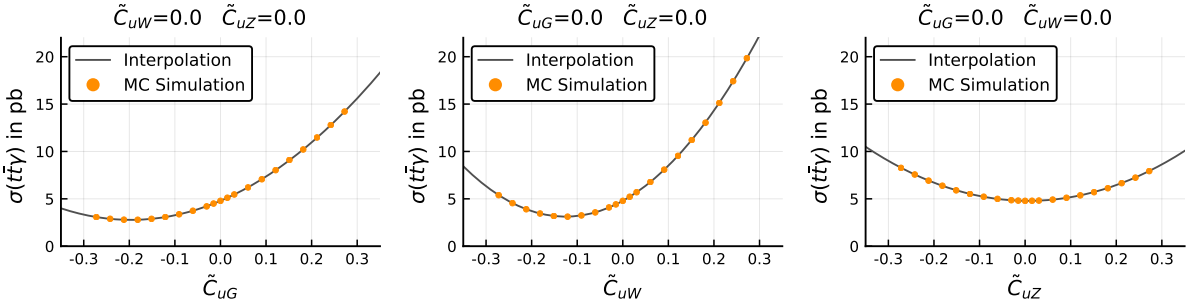

Figure 7.5: Sampling points and result of the interpolation for the total $t\bar{t}\gamma$ cross section, shown as slices of the phase space where only one of the Wilson coefficients is varied at a time, while the others are fixed to zero.

A very good agreement of the interpolation with the sampling points is observed, with the standard deviation of the relative deviations at all sampling points being only 0.2 %. As the measurement in Ref. [201] is performed in a fiducial region and no unfolded results are given, fiducial acceptances are calculated for the simulated cross sections. To determine them, parton showering is applied to the MC samples using `PYTHIA8` [217]. A particle-level event selection is performed at each sampling point using the `MadAnalysis` framework [218–221] to recover the same fiducial region as in the corresponding ATLAS measurement. The clustering of particle jets is conducted using the anti-k_t algorithm [222] with a radius parameter of $R = 0.4$ using `FastJet` [223]. The fiducial acceptances obtained for the SM-like sampling point agree well with those of Ref. [201], having the same value in the case of the dilepton channel and only a deviation of 3 % for the single-lepton channel. It has been checked explicitly that performing only a parton-level simulation and applying the fiducial cuts at this level is not a suitable approximation, as it leads to deviations of the LO SM predictions by about 50 % for the single-lepton channel and by about 25 % for the dilepton channel.

The values of the fiducial acceptances, A, obtained at the sampling points are parametrized according to

$$A = \frac{A^{\mathrm{SM}}\sigma^{\mathrm{SM}} + \sum_i \tilde{C}_i A_i^{\mathrm{interf.}} \sigma_i^{\mathrm{interf.}} + \sum_{i\leq j} \tilde{C}_i \tilde{C}_j A_{ij}^{\mathrm{BSM}} \sigma_{ij}^{\mathrm{BSM}}}{\sigma^{\mathrm{SM}} + \sum_i \tilde{C}_i \sigma_i^{\mathrm{interf.}} + \sum_{i\leq j} \tilde{C}_i \tilde{C}_j \sigma_{ij}^{\mathrm{BSM}}} \,. \tag{7.5}$$

Here, the denominator is the parametrization of the total cross section, determined before and shown in Fig. 7.5. The unknown parameters A_i and A_{ij} of the fiducial acceptances are determined using a least-squares fit to Eq. (7.5). The obtained parametrization of the fiducial acceptances and the corresponding sampling points for the single-lepton channel are shown in Fig. 7.6. The plots of the dilepton channel can be found in Fig. D.1 of App. D.

The desired dependence of the fiducial cross sections on the Wilson coefficients $\tilde{C}_{uB}$, $\tilde{C}_{uG}$, and $\tilde{C}_{uW}$ is finally obtained by multiplying the interpolation of the total cross section with the interpolation of the fiducial acceptances. NLO SM calculations of the fiducial cross sections are taken into account by applying a k-factor, shifting the parametrization so that the SM contributions correspond to the according values of the NLO predictions. The resulting

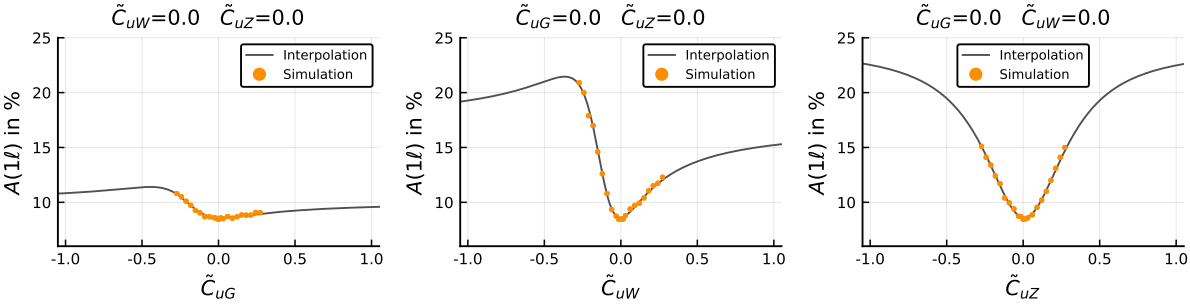

Figure 7.6: Sampling points and result of the interpolation for the fiducial acceptance of the single-lepton channel $A(1\ell)$. Shown are slices of the phase space where only one of the Wilson coefficients is varied at a time, while the others are fixed to zero.

parametrizations of the fiducial cross sections as a function of the coefficients $\tilde{C}_{uB}$, $\tilde{C}_{uG}$, and $\tilde{C}_{uW}$ are shown in Fig. 7.7 for the single-lepton (left) and dilepton (right) cases. Again, only one of the Wilson coefficients is varied, and the others are fixed to zero. Also indicated are the measurements of Ref. [201] as well as the associated total uncertainty, which includes experimental and theory uncertainties. The parametrizations of the single-lepton and dilepton channels yield very similar sensitivities to the three Wilson coefficients considered.

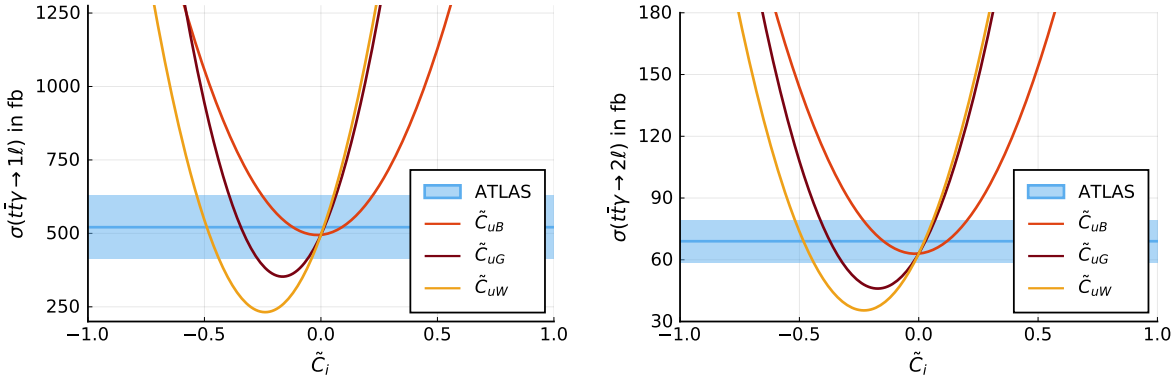

Figure 7.7: Parametrization of the fiducial $t\bar{t}\gamma$ cross sections for the single-lepton channel (left) and the dilepton channel (right) when only one of the Wilson coefficients is varied at a time, while the others are set to zero. The corresponding ATLAS measurement and the combined experimental and theory uncertainties are indicated by the blue line and band.

7.4 Constraining the SMEFT Wilson coefficients

The dependences of the $t\bar{t}\gamma$ cross section and of the $\bar{B} \to X_s\gamma$ branching ratio on the SMEFT Wilson coefficients shown in Fig. 7.7 and Fig. 7.4, respectively, are implemented in EFTfitter.jl, and the measurements of Tab. 7.1 are inserted together with the corresponding experimental and theory uncertainties. Since no information about correlations between the measurements is available, they are initially assumed to be uncorrelated. The influence of potential correlations on the results is investigated in a subsequent step. A uniform prior distribution is assumed for

all three Wilson coefficients within the range $-1 \leq \tilde{C}_i \leq 1$, since larger values of the Wilson coefficients would lead to a breakdown of the EFT expansion. As the parameter space is low-dimensional, the Sobol sampler of BAT.jl is applied to explore the posterior distribution.

When performing a SMEFT fit of all three coefficients using only the measurement of the $\bar{B} \to X_s\gamma$ branching fraction, only constraints on $\tilde{C}_{uB}$ can be derived, while $\tilde{C}_{uG}$ and $\tilde{C}_{uW}$ remain unconstrained. The marginalized posterior distribution of $\tilde{C}_{uB}$ is visualized in the left plot of Fig. 7.8. The smallest intervals containing 90 % of the posterior probability are highlighted.

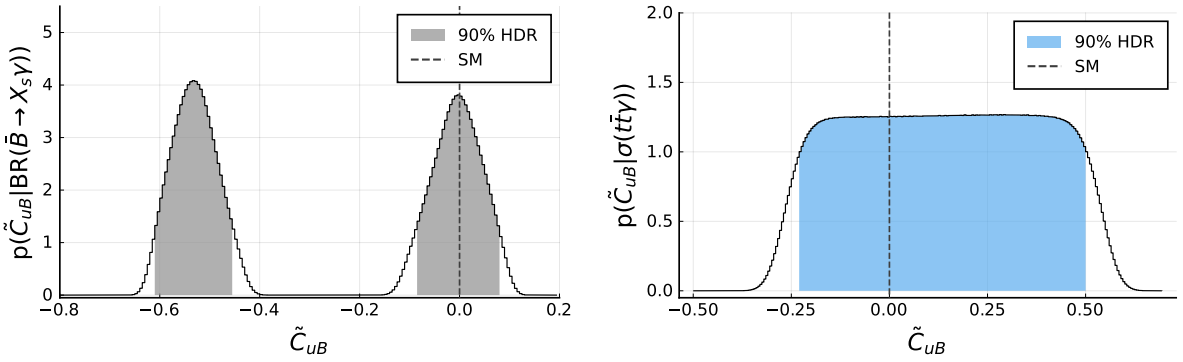

Figure 7.8: Marginalized posterior distribution of $\tilde{C}_{uB}$ obtained when using only the $\mathrm{BR}(\bar{B} \to X_s\gamma)$ measurement (left) and when using only the $t\bar{t}\gamma$ cross section measurement (right). The colored areas highlight the 90 % HDRs, and the dashed lines indicate the SM values.

For $\tilde{C}_{uB}$, two regions of the parameter space are in agreement with the $\mathrm{BR}(\bar{B} \to X_s\gamma)$ measurement. One region is around the SM value, and the other one is at negative values around $\tilde{C}_{uB} \approx -0.5$. This result is expected when considering the quadratic shape of the sensitivity of $\mathrm{BR}(\bar{B} \to X_s\gamma)$ to $\tilde{C}_{uB}$ in Fig. 7.4. Without further input, however, neither of the two allowed regions can be excluded. Indeed, this ambiguity is well-known in B physics and can be resolved using measurements of semileptonic $b \to s\ell^+\ell^-$ decays, which allows excluding the solution around $\tilde{C}_{uB} \approx -0.5$ [224, 225]. Such observables will be included in the studies of Chap. 8. Here, it is investigated whether a combination with top-quark observables can also resolve this ambiguity.

When using only the measurements of the $t\bar{t}\gamma$ production cross section, constraints on all three coefficients can be derived. The corresponding marginalized posterior distribution of $\tilde{C}_{uB}$ is shown in the right plot of Fig. 7.8. The distributions of $\tilde{C}_{uG}$ and $\tilde{C}_{uW}$ have similar shapes and are shown in Fig. D.2 of App. D. The constraints obtained from the $t\bar{t}\gamma$ cross sections are of a similar magnitude for all three Wilson coefficients with the widths of their 90 % HDRs being of the order $\mathcal{O}(1)$.

When finally combining the top-quark and B physics measurements in a joint fit, the constraints on the Wilson coefficients improve significantly, in particular for $\tilde{C}_{uB}$. This can be observed in Fig. 7.9, where the one- and two-dimensional 90 % HDRs are shown for all three Wilson coefficients. The HDRs are given for the fit using only the $\mathrm{BR}(\bar{B} \to X_s\gamma)$ measurements (gray), for the fit using only the $t\bar{t}\gamma$ cross section measurements (light blue), and for the

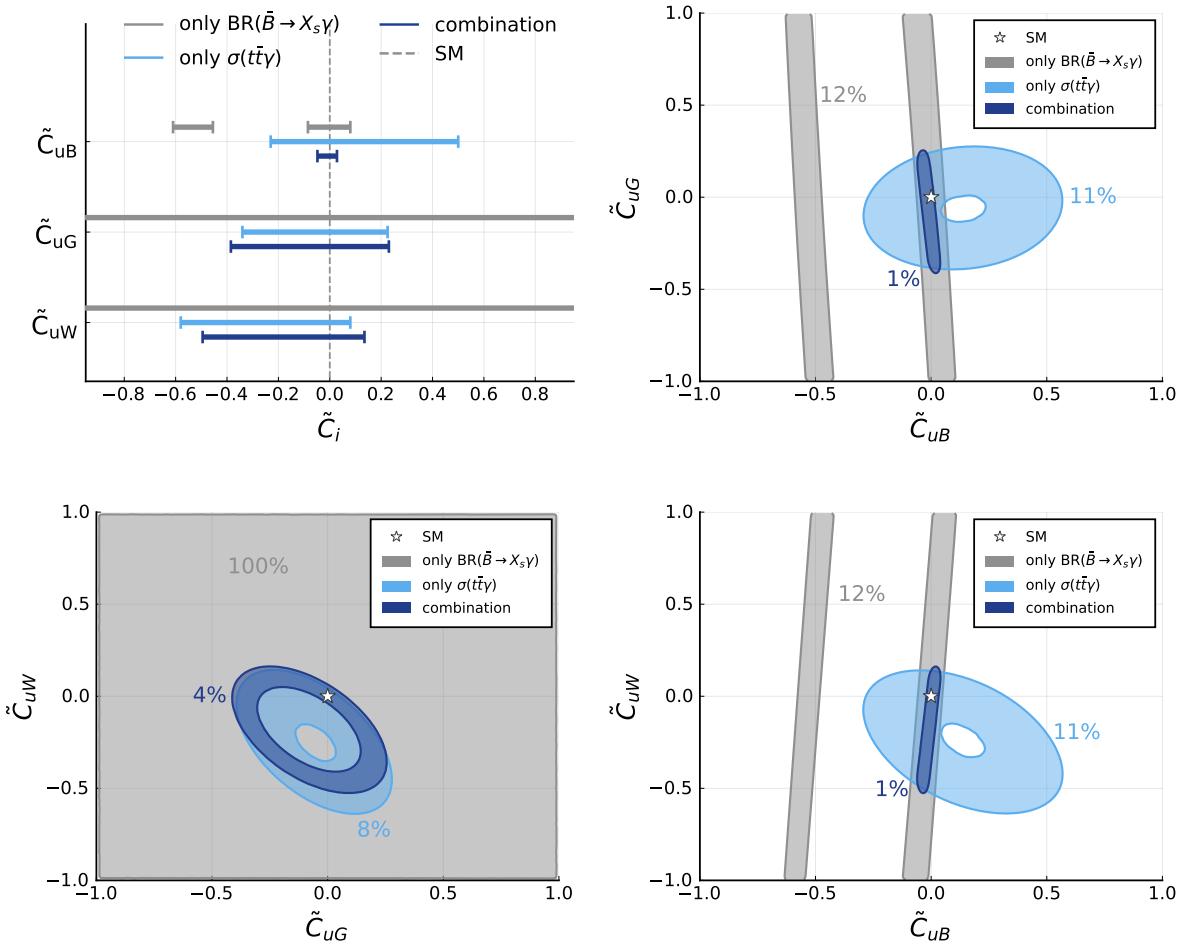

Figure 7.9: One- and two-dimensional marginalized 90 % HDRs obtained in fits to only BR($\bar{B} \to X_s\gamma$) (gray), to only $\sigma(t\bar{t}\gamma)$ (light blue), and to their combination (dark blue). The SM value is indicated. The numbers in the two-dimensional plots indicate the relative sizes of the corresponding areas in the two-dimensional parameter space spanned by $-1 \leq \tilde{C}_i \leq 1$.

combined fit to all measurements (dark blue). In both upper plots as well as in the lower right plot of Fig. 7.9, it can be observed that including the $t\bar{t}\gamma$ cross sections allows to resolve the ambiguity in the constraints on $\tilde{C}_{uB}$ obtained in the B-only fit. While in the one-dimensional HDRs of the upper left plot no large improvement of the constraints on $\tilde{C}_{uG}$ and $\tilde{C}_{uW}$ is visible in the combination compared to the top-only fit, when considering the two-dimensional marginal distributions, it can be observed how the allowed parameter space is tightened in the combined fit. In particular in the two plots on the right-hand side of Fig. 7.9, showing the two-dimensional marginal distributions of $\tilde{C}_{uB}$ vs. $\tilde{C}_{uG}$ and $\tilde{C}_{uB}$ vs. $\tilde{C}_{uW}$, it is observed that the sensitivity of the top-quark and B observables to these Wilson coefficients are rather orthogonal, which leads to the significantly improved constraints in their combination. To quantify this improvement of the constraints, the sizes of the two-dimensional HDRs depicted in Fig. 7.9 are compared. In both plots on the right-hand side, the constraints obtained in the B-only fit cover about 12 % of the allowed two-dimensional parameter spaces spanned by $\tilde{C}_{uB}, \tilde{C}_{uG} \in [-1, 1]$ and $\tilde{C}_{uB}, \tilde{C}_{uW} \in [-1, 1]$, respectively. The constraints from the fit using

only the $t\bar{t}\gamma$ cross sections are of a similar size for both projections and cover about 11 % of the allowed space. Due to the complementarity of the constraints from top-quark and B physics, the allowed region in the combined fit is reduced by more than an order of magnitude, with the corresponding area of the combination taking only about 1 % of the allowed parameter space. The two-dimensional marginal distribution of $\tilde{C}_{uG}$ vs. $\tilde{C}_{uW}$ in the lower-left plot of Fig. 7.9 does not depend on $\tilde{C}_{uB}$ and is therefore not directly constrained by the $\mathrm{BR}(\bar{B} \to X_s\gamma)$ measurements. However, also in this projection, an improvement of the 90 % HDR is observed in the combined fit, with the allowed area being reduced by a factor of about 1.9 when compared to the top-only fit. This is the result of a reduction of the allowed three-dimensional parameter space, which also translates into the two-dimensional marginal distributions.

As only three Wilson coefficients are considered in this fit, the full three-dimensional posterior distributions can be visualized. In Fig. 7.10, two different views of the three-dimensional posterior distributions obtained in the fits to the individual data sets and in the combined fit are shown. In this representation, it is clearly visible that the fit using only $\mathrm{BR}(\bar{B} \to X_s\gamma)$ allows for two separated plane-shaped 90 % HDRs and constrains only $\tilde{C}_{uB}$. The measurements of the $t\bar{t}\gamma$ cross section constrain all three Wilson coefficients to a certain degree, and the corresponding posterior distribution has the shape of an ellipsoid shell. It is visible that the ring-shaped posterior distribution of the combined fit is the result of one of the planes of the B-only constraints intersecting the ellipsoid of the top-only fit. When considering the three-dimensional 90 % HDVs, determined as described in Sec. 5.4, the reduction of the allowed parameter space becomes even more apparent than when considering only the two-dimensional marginal distributions. The posterior distribution of the B-only fit has a 90 % HDV of about 0.333, while that of the top-only fit has a 90 % HDV of 0.138. In the combined fit, the 90 % HDV is more than an order of magnitude smaller and is only 0.009.

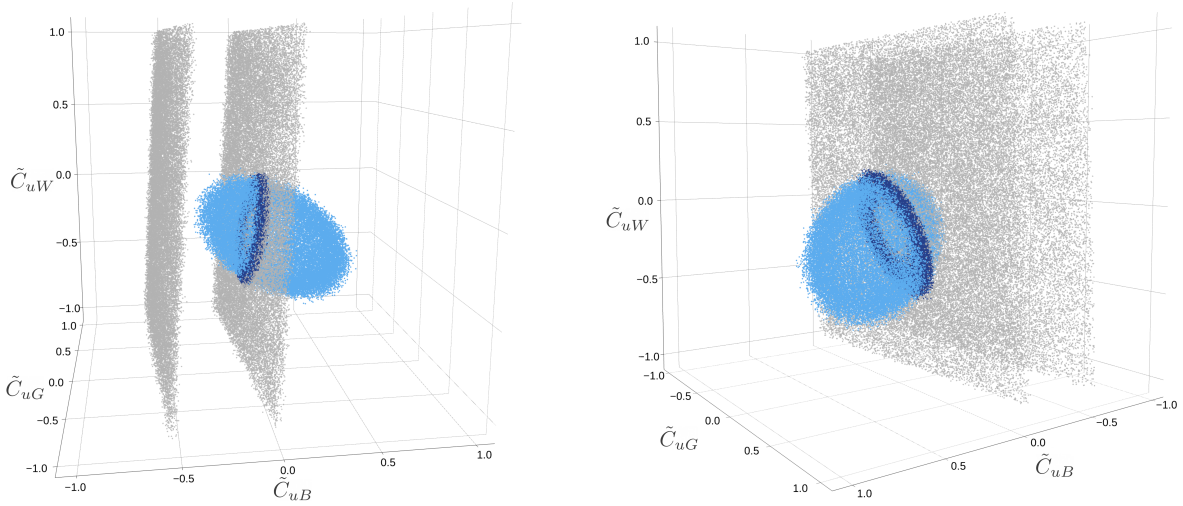

Figure 7.10: Two views of the three-dimensional posterior distributions obtained in fits to only $\mathrm{BR}(\bar{B} \to X_s\gamma)$ (gray), to only $\sigma(t\bar{t}\gamma)$ (light blue), and to their combination (dark blue).

In the fits discussed above, the measurements were assumed to be completely uncorrelated, regarding both the experimental and the theory uncertainties. While this assumption is well-justified for the correlations between the B physics and top-quark measurements, correlations between the two considered channels of the $t\bar{t}\gamma$ cross section measurements might be present and should be considered, as discussed in Chap. 6. The actual values of the respective correlation coefficients, however, are not known. Therefore, the effects of non-zero correlations for the experimental and theory uncertainties between the single-lepton and dilepton channels are investigated by varying the values of the correlation coefficients. For the systematic uncertainties, the correlation coefficient is varied between -0.9 and 0.9. For the theory uncertainties, only positive correlations are considered, varying the corresponding correlation coefficient between 0 and 0.9. When performing the combined fit with varying correlation coefficients, only minor changes in the sizes of the HDRs of $\tilde{C}_{uB}$ vs. $\tilde{C}_{uG}$ and $\tilde{C}_{uB}$ vs. $\tilde{C}_{uW}$ by about 4 % are observed. No changes in the overall shape or position of the posterior distributions are visible. As the marginal distribution of $\tilde{C}_{uG}$ vs. $\tilde{C}_{uW}$ is mainly constrained by the $t\bar{t}\gamma$ cross section measurement, slightly larger changes can be observed when varying the correlation coefficients. Again, the general shape and position of the constraints are not affected, and only the width of the ring changes, with the 90 % HDA increasing by up to 30 % compared to the uncorrelated scenario. However, these differences do not impact the essential findings discussed above. Since the main objective of this study is to demonstrate the feasibility of combining top-quark and B physics data, the resulting benefits shown above are not affected by potential correlations between the two channels of the $t\bar{t}\gamma$ cross section measurement.

7.5 Summary

It was demonstrated that observables from the B physics sector can be employed to constrain the Wilson coefficients of dimension-six SMEFT operators affecting top-quark couplings. The steps necessary for translating WET coefficients into SMEFT coefficients were outlined and allow determining BSM contributions to B observables as a function of SMEFT Wilson coefficients at the top-quark scale. In this proof-of-concept study with only two observables, it was revealed that a combination of top-quark and B measurements can improve the constraints on top-quark Wilson coefficients significantly. Due to the observed complementarity of constraints derived from the individual observables, including B physics processes from a lower energy scale into global fits of SMEFT Wilson coefficients of the top-quark sector seems to be a promising approach for enhancing SMEFT fits. In the following chapter, the observed synergies between top-quark and flavor physics observables are further explored, and their potential in future scenarios is investigated.

8 Synergies of top-quark and B physics in current and future SMEFT fits

In the previous chapter, it was demonstrated that the combination of observables from top-quark and B physics can significantly improve the constraints on SMEFT Wilson coefficients. In this chapter, this approach is extended by including more observables and coefficients in a combined fit. In addition, the impact of potential future data is examined by including projections for measurements expected in the near and far future. Particular focus is placed on the challenges of sampling the resulting multimodal posterior distributions. This chapter is organized as follows: In Sec. 8.1, the considered observables are introduced, and an outline of the following studies is given. In Sec. 8.2, the SMEFT coefficients to be constrained are stated, and the sensitivities of the different observables are discussed. Fits of the Wilson coefficients using current measurements are performed in Sec. 8.3.3, and the synergies between top-quark and B physics observables are highlighted. Potential improvements of the constraints when considering future measurements from HL-LHC, Belle II, and CLIC are explored in Sec. 8.4. In Sec. 8.5, it is demonstrated that only the nested sampling algorithms of BAT.jl are suited for sampling the multimodal posterior distributions obtained in the combined fits. A brief summary of the results is given in Sec. 8.6.

8.1 Extending combinations of top-quark and B physics observables

The steps necessary for combining observables from top-quark and flavor physics in a joint fit of SMEFT Wilson coefficients were outlined in Chap. 7. In the following, this approach is extended, and further observables and measurements from top-quark and B physics are considered, allowing to constrain a significantly larger set of Wilson coefficients.

From the top-quark sector, the measurements of the fiducial $t\bar{t}\gamma$ production cross section are again included. In addition, measurements of $t\bar{t}$ production and $t\bar{t}Z$ production cross sections are considered. Also, top-quark decay observables, namely the W boson helicity fractions and the top-quark decay width, are included. From B physics, the $\bar{B} \to X_s\gamma$ branching ratio as an observable of $b \to s\gamma$ transitions is again taken into account. In addition, observables of $b \to s\ell^+\ell^-$ transitions, such as (differential) branching ratios and multiple angular observables, are considered. Furthermore, cross sections, asymmetries, and ratios of $Z \to b\bar{b}$ decays are included in the fit for improving the sensitivity to certain coefficients, as will be discussed in the next section. In further steps of the study, projections for measurements expected from future experiments are considered. In a first step, a near-future scenario is investigated in which measurements of the same observables but with improved precision are expected to be obtained by the HL-LHC experiments and Belle II. Additionally, in this scenario, measurements of $b \to s\nu\bar{\nu}$ transitions are assumed to be achieved by the Belle II experiment. In a next step, a far-future scenario is considered, in which a CLIC-like lepton collider is expected to provide

measurements of $e^+e^- \to t\bar{t}$ processes. The specific sensitivities of the considered processes to the Wilson coefficients of dimension-six SMEFT operators are discussed in the next section.

8.2 Sensitivities to the Wilson coefficients

While in the simple example of Chap. 7 only three SMEFT Wilson coefficients were considered, the additionally included observables allow constraining more parameters. Therefore, the Wilson coefficients of up to eleven dimension-six operators are considered here. These are those of the bosonic operators

$$\begin{aligned} O_{uB} &= (\bar{q}_L \sigma^{\mu\nu} u_R)\, \tilde{\varphi} B_{\mu\nu}\,, & O_{uG} &= \left(\bar{q}_L \sigma^{\mu\nu} T^A u_R\right) \tilde{\varphi} G^A_{\mu\nu}\,, \\ O_{uW} &= \left(\bar{q}_L \sigma^{\mu\nu} \tau^I u_R\right) \tilde{\varphi} W^I_{\mu\nu}\,, & O^{(1)}_{\varphi q} &= \left(\varphi^\dagger i \overleftrightarrow{D}_\mu \varphi\right) (\bar{q}_L \gamma^\mu q_L)\,, \\ O_{\varphi u} &= \left(\varphi^\dagger i \overleftrightarrow{D}_\mu \varphi\right) (\bar{u}_R \gamma^\mu u_R)\,, & O^{(3)}_{\varphi q} &= \left(\varphi^\dagger i \overleftrightarrow{D}^I_\mu \varphi\right) \left(\bar{q}_L \tau^I \gamma^\mu q_L\right)\,, \end{aligned} \tag{8.1}$$

and those of the semileptonic four-fermion operators

$$\begin{aligned} O^{(1)}_{lq} &= \left(\bar{l}_L \gamma_\mu l_L\right) (\bar{q}_L \gamma^\mu q_L)\,, & O^{(3)}_{lq} &= \left(\bar{l}_L \gamma_\mu \tau^I l_L\right) \left(\bar{q}_L \gamma^\mu \tau^I q_L\right)\,, \\ O_{qe} &= (\bar{q}_L \gamma_\mu q_L)\, (\bar{e}_R \gamma^\mu e_R)\,, & O_{eu} &= (\bar{e}_R \gamma_\mu e_R)\, (\bar{u}_R \gamma^\mu u_R)\,, \\ O_{lu} &= \left(\bar{l}_L \gamma_\mu l_L\right) (\bar{u}_R \gamma^\mu u_R)\,. \end{aligned} \tag{8.2}$$

Since the dominant BSM contributions at order $\mathcal{O}(\Lambda^{-2})$ are proportional to the real parts of the operators, again, only real-valued Wilson coefficients are taken into account. Details on the definitions of these operators and considered contributions from down-type quarks to the corresponding coefficients are given in Ref. [5]. As pointed out before, different observables yield different sensitivities to the Wilson coefficients. Some observables, however, are only sensitive to certain linear combinations of the dimension-six operators, as was also the case in the example of Sec. 6.2. Here, the linear combinations

$$\tilde{C}^{\pm}_{\varphi q} = \tilde{C}^{(1)}_{\varphi q} \pm \tilde{C}^{(3)}_{\varphi q} \qquad \text{and} \qquad \tilde{C}^{\pm}_{lq} = \tilde{C}^{(1)}_{lq} \pm \tilde{C}^{(3)}_{lq} \tag{8.3}$$

are constrained by some of the observables.

The dominant sensitivities of the observables from the different physics sectors to the SMEFT coefficients considered here are indicated in Fig. 8.1. The specific observables included in the studies, their sensitivities to the Wilson coefficients, and the computation of corresponding BSM contributions are discussed in detail in the following. A summary of the observables and their dominant and subdominant sensitivities to the coefficients is given in Tab. 8.1.

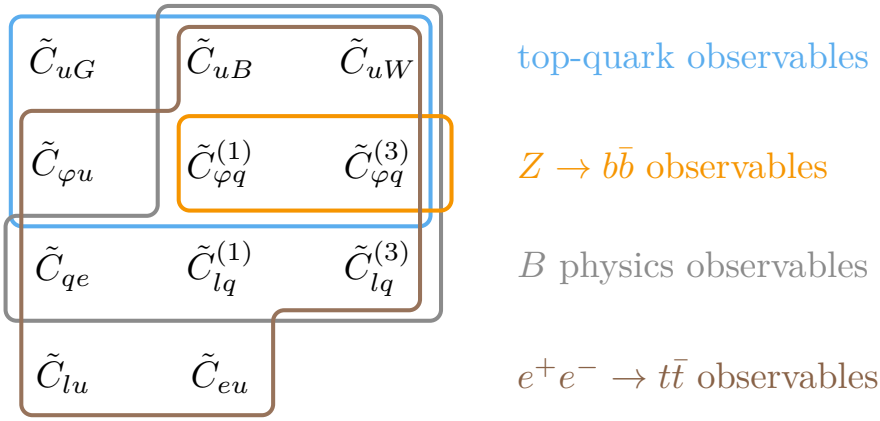

Figure 8.1: Dominant sensitivities of the different types of observables to the SMEFT coefficients considered here. Adapted from Ref. [5].

8.2.1 Sensitivities of top-quark observables

As shown in Chap. 7, the leading BSM contributions to the $t\bar{t}\gamma$ production cross section come from the coefficients $\tilde{C}_{uB}$, $\tilde{C}_{uG}$ and $\tilde{C}_{uW}$. Here, also the cross section of $t\bar{t}$ production is included to improve the sensitivity to $\tilde{C}_{uG}$. The top-quark decay observables, i.e., the top-quark decay width Γ_t and the W boson helicity fractions F_0 and F_L, allow constraining the coefficients $\tilde{C}_{uW}$ and $\tilde{C}^{(3)}_{\varphi q}$, as already shown in Chap. 6. Furthermore, the cross section of $t\bar{t}Z$ production is considered, which depends on $\tilde{C}_{uB}$, $\tilde{C}_{uG}$ and $\tilde{C}_{uW}$ as well, but is also sensitive to BSM contributions from $\tilde{C}_{\varphi u}$ and from the linear combination $\tilde{C}^{-}_{\varphi q}$. By combining measurements of the top-quark decay width, sensitive to $\tilde{C}^{(3)}_{\varphi q}$, with measurements of the $t\bar{t}Z$ cross section, sensitive to $\tilde{C}^{-}_{\varphi q}$, it is thus possible to probe $\tilde{C}^{(1)}_{\varphi q}$ and $\tilde{C}^{(3)}_{\varphi q}$ individually, according to Eq. (8.3). Therefore, the top-quark observables considered here allow constraining a total of six coefficients of two-fermion operators, namely

$$\tilde{C}_{uB}\,, \quad \tilde{C}_{uG}\,, \quad \tilde{C}_{uW}\,, \quad \tilde{C}^{(1)}_{\varphi q}\,, \quad \tilde{C}^{(3)}_{\varphi q}\,, \quad \tilde{C}_{\varphi u}\,. \tag{8.4}$$

The computation of BSM contributions to the top-quark cross sections as a function of these Wilson coefficients is again performed using MadGraph5_aMC@NLO and the `dim6top_LO` UFO model with a subsequent interpolation according to Eq. (4.8). For the fiducial cross sections of $t\bar{t}\gamma$ production, the model derived in Sec. 7.3.2 is employed. Similar to Chap. 6, the dependence of the W boson helicity fractions on $\tilde{C}_{uW}$ is computed according to Ref. [183], with the addition that also quadratic contributions are included. For the top-quark decay width, the BSM contributions are calculated according to Ref. [18], also including quadratic contributions.

8.2.2 Sensitivities of $Z \to b\bar{b}$ observables

Observables of $Z \to b\bar{b}$ processes in e^+e^- collisions are included here to tighten the constraints on $\tilde{C}^{(1)}_{\varphi q}$ and $\tilde{C}^{(3)}_{\varphi q}$. Specifically, the forward-backward asymmetry $A^{0,b}_{\mathrm{FB}}$, the ratio of partial widths for $Z \to f\bar{f}$ decays $R_b = \Gamma_{b\bar{b}}/\Gamma_{\mathrm{had.}}$, and the hadronic cross section $\sigma_{\mathrm{had.}}$ are considered

here. These observables are sensitive to the linear combination $\tilde{C}^+_{\varphi q}$. As this is the linear combination with the opposite sign as in $\tilde{C}^-_{\varphi q}$, which is constrained by the $t\bar{t}Z$ cross section, improved constraints on $\tilde{C}^{(1)}_{\varphi q}$ and $\tilde{C}^{(3)}_{\varphi q}$ are expected when combining the $Z \to b\bar{b}$ observables with the top-quark observables.

The BSM contributions to $Z \to b\bar{b}$ processes as a function of $\tilde{C}^+_{\varphi q}$ are also computed using MadGraph5_aMC@NLO and the `dim6top_LO` UFO model. For the hadronic cross section, a quadratic interpolation to Eq. (4.8) is performed. For the forward-backward asymmetry

$$A_{\mathrm{FB}} = \frac{\sigma_{\mathrm{FB}}}{\sigma}\,, \quad \text{with} \quad \sigma_{\mathrm{FB}} = \int_{-1}^{1} \mathrm{sign}(\cos\theta)\frac{\mathrm{d}\sigma}{\mathrm{d}\cos\theta}\mathrm{d}\cos\theta\,, \tag{8.5}$$

where θ is the angle between the momenta of the bottom quark and the positron in the center-of-mass frame, the BSM contributions are interpolated according to Eq. (4.8) for both the numerator and denominator. Similarly, for R_b, the BSM contributions are parametrized using quadratic interpolations of the partial widths according to

$$\Gamma = \Gamma^{\mathrm{SM}} + \tilde{C}^+_{\varphi q}\Gamma^{\mathrm{int.}} + (\tilde{C}^+_{\varphi q})^2\Gamma^{\mathrm{BSM}} \tag{8.6}$$

for both the numerator and denominator.

8.2.3 Sensitivities of B physics observables

In addition to the contributions from $\tilde{C}_{uB}$, $\tilde{C}_{uG}$, and $\tilde{C}_{uW}$, studied in the previous chapter, for $b \to s\gamma$ transitions now also contributions from $\tilde{C}^{(3)}_{\varphi q}$ are considered. Next to the branching ratio of $\bar{B} \to X_s\gamma$, the branching ratios of $B^0 \to K^*\gamma$ and $B^+ \to K^{*+}\gamma$ are included as observables of such transitions. Several observables of $b \to s\ell^+\ell^-$ transitions are taken into account, such as (differential) branching ratios, angular observables P_i, and forward-backward asymmetries A_{FB}. These quantities are included for various processes. In particular, these are different types of $B \to K\mu\mu$ decays, but also $B_s \to \mu^+\mu^-$, $B_s \to \phi\mu^+\mu^-$, and $\Lambda_b \to \Lambda\mu^+\mu^-$. A detailed list of the specific observables included is given in Tab. 8.3. The $b \to s\ell^+\ell^-$ processes are sensitive to $\tilde{C}_{uB}$, $\tilde{C}_{uG}$, $\tilde{C}_{uW}$, and $\tilde{C}^{(3)}_{\varphi q}$ or $\tilde{C}^+_{\varphi q}$. In addition, they can also receive contributions from four-fermion operators, namely from $\tilde{C}_{qe}$ and from the linear combination $\tilde{C}^+_{lq}$. Furthermore, the mass difference ΔM_s of $B_s - \bar{B}_s$ mixing is included, which is sensitive to $\tilde{C}_{uW}$, $\tilde{C}_{uG}$, and $\tilde{C}^{(3)}_{\varphi q}$. Therefore, the B physics observables considered here provide sensitivity to the eight coefficients

$$\tilde{C}_{uB}\,, \quad \tilde{C}_{uG}\,, \quad \tilde{C}_{uW}\,, \quad \tilde{C}^{(1)}_{\varphi q}\,, \quad \tilde{C}^{(3)}_{\varphi q}\,, \quad \tilde{C}_{\varphi u} \quad \tilde{C}_{qe}\,, \quad \tilde{C}^+_{lq}\,. \tag{8.7}$$

The BSM contributions to $b \to s\gamma$ and $b \to s\ell^+\ell^-$ transitions are computed in LO QCD using the `flavio` and `wilson` packages. The dependence on the SMEFT coefficients at the top-quark scale is again determined by performing the matching and running procedure as described in Sec. 7.1.2. For the mass difference of $B_s - \bar{B}_s$ mixing, the BSM contributions in terms of WET are determined according to Ref. [226], as detailed in Ref. [5].

Table 8.1: Overview of the considered observables and their sensitivities to the SMEFT coefficients. Coefficients without brackets contribute at tree level. Coefficients in square brackets, $[\tilde{C}_i]$, contribute only at one-loop level to B physics observables. Coefficients written in braces, $\{\tilde{C}_i\}$, are induced by SMEFT and WET running only. Coefficients marked with asterisks, $\tilde{C}_i^{(*)}$ and $\tilde{C}_i^{(**)}$, receive contributions at one-loop level that change their tree-level definitions, as given in App. D.1. Observables and corresponding Wilson coefficients written in gray are considered only in the future projections and are introduced in Sec. 8.4.1 and Sec. 8.4.2. Adapted from Ref. [5].

Process	Observable	Two-fermion	Four-fermion
$pp \to t\bar{t}$	σ^{inc}	$\tilde{C}_{uG}$	-
$pp \to t\bar{t}\gamma$	σ^{fid}	$\tilde{C}_{uB}$, $\tilde{C}_{uW}$, $\tilde{C}_{uG}$	-
$pp \to t\bar{t}Z$	σ^{inc}	$\tilde{C}_{uB}$, $\tilde{C}_{uW}$, $\tilde{C}_{uG}$, $\tilde{C}^{-}_{\varphi q}$, $\tilde{C}_{\varphi u}$	-
$t \to bW$	$F_{0,L}$	$\tilde{C}_{uW}$	-
Top decay	Γ_t	$\tilde{C}^{(3)}_{\varphi q}$, $\tilde{C}_{uW}$	-
$Z \to b\bar{b}$	A^b_{FB}, R_b, σ_{had}	$\tilde{C}^{+}_{\varphi q}$	-
$b \to s\gamma$	BR	$\left[\tilde{C}_{uB}\right]$, $\left[\tilde{C}_{uW}\right]$, $\left\{\tilde{C}_{uG}\right\}$, $\left[\tilde{C}^{(3)}_{\varphi q}\right]$	-
$b \to s\ell^+\ell^-$	BR, A_{FB}, $P_i^{(\prime)}$, S_i, F_L, $\text{dBR}/\text{d}q^2$	$\left[\tilde{C}_{uB}\right]$, $\left[\tilde{C}_{uW}\right]$, $\left\{\tilde{C}_{uG}\right\}$, $\tilde{C}^{+(*)}_{\varphi q}$, $\left[\tilde{C}^{(3)}_{\varphi q}\right]$	$\tilde{C}^{+(*)}_{lq}$, $\tilde{C}^{(*)}_{qe}$
Mixing	ΔM_s	$\left[\tilde{C}_{uW}\right]$, $\left\{\tilde{C}_{uG}\right\}$, $\left[\tilde{C}^{(3)}_{\varphi q}\right]$	-
$b \to s\nu\bar{\nu}$	BR	$\tilde{C}^{+(**)}_{\varphi q}$	$\tilde{C}^{-(*)}_{lq}$
$e^+e^- \to t\bar{t}$	σ, A_{FB}	$\tilde{C}_{uB}$, $\tilde{C}_{uW}$, $\left\{\tilde{C}_{uG}\right\}$, $\tilde{C}^{-}_{\varphi q}$, $\tilde{C}_{\varphi u}$	$\tilde{C}_{eu}$, $\tilde{C}_{qe}$, $\tilde{C}_{lu}$, $\tilde{C}^{-}_{lq}$

8.3 Constraining SMEFT coefficients with current data

The observables introduced above and the parametrizations of their dependences on the SMEFT coefficients are inserted into EFTfitter.jl. Analogously to Chap. 7, constraints on the coefficients are derived using both the individual and the combined data sets to demonstrate the gain in sensitivity when combining observables from different physics sectors. The measurements included in the SMEFT fits are introduced below. The specific observables and corresponding measurements are listed in Tabs. 8.2, 8.3, and 8.4. Experimental uncertainties, consisting of statistical and systematic contributions, as well as theory uncertainties on the SM predictions, are taken into account in all fits. As demonstrated in Chap. 6, unknown correlations between the uncertainties of measurements can have a significant impact on the results of their combination. All available information on correlations, e.g., for the measurements of W boson helicity fractions, is included. To minimize the effects of unknown correlation coefficients, in the following only the most precise measurement of each observable is considered, similar to Refs. [17, 227]. Wherever available, averaged values from the HFLAV group are used for the B physics observables. In the case of distributions, only the bin with the smallest uncertainties is considered, as particularly bin-to-bin correlations can have significant impacts on the results. Values of correlations within the same q^2 bins are available and taken into

account. Given these measures, the simplifying assumption of the remaining measurements being uncorrelated is well justified, especially considering that the primary purpose of the following studies is to point out the synergies of observables from different energy scales rather than to provide a fit including all data available.

8.3.1 Constraints from top-quark measurements

A fit considering only the top-quark observables introduced in Sec. 8.2.1 is performed to constrain the six coefficients of Eq. (8.4). For the $t\bar{t}\gamma$ cross sections, again the fiducial ATLAS measurements of Ref. [201] are included, using both the single-lepton and dilepton channels. For the $t\bar{t}Z$ cross section, a measurement by the CMS experiment at 13 TeV [11] is used, and for the $t\bar{t}$ cross section, a measurement by ATLAS at 13 TeV [228] is employed. Measured values of the W boson helicity fractions and the top-quark decay width are again taken from the respective ATLAS measurements at 8 TeV in Ref. [163] and Ref. [162]. The considered top-quark observables and the corresponding references for the measurements and SM predictions are summarized in Tab. 8.2.

Table 8.2: Considered observables of top-quark processes and references of the included measurements and SM predictions.

Process	Observable	$\sqrt{s}$	Int. luminosity	Experiment	Ref.	SM Ref.
$t\bar{t}\gamma$	$\sigma^{\text{fid}}(t\bar{t}\gamma, 1\ell)$, $\sigma^{\text{fid}}(t\bar{t}\gamma, 2\ell)$	13 TeV	36.1 fb^{-1}	ATLAS	[201]	[201, 206]
$t\bar{t}Z$	$\sigma^{\text{inc}}(t\bar{t}Z)$	13 TeV	77.5 fb^{-1}	CMS	[11]	[229–231]
$t\bar{t}$	$\sigma^{\text{inc}}(t\bar{t})$	13 TeV	36.1 fb^{-1}	ATLAS	[228]	[232]
	F_0 , F_L	8 TeV	20.2 fb^{-1}	ATLAS	[163]	[177]
	Γ_t	8 TeV	20.2 fb^{-1}	ATLAS	[162]	[184]

For the Bayesian fit with EFTfitter.jl, a uniform prior distribution in the range $-1 \leq \tilde{C}_i \leq 1$ is assumed for all six Wilson coefficients. Again, values $|\tilde{C}_i| > 1$ are not reasonable, as for such, the EFT expansion breaks down. The Metropolis–Hastings algorithm is employed for sampling the six-dimensional posterior distribution. The one-dimensional marginalized intervals containing 90 % of the posterior probability obtained in this fit to only the top-quark data of Tab. 8.2 are shown in the left plot of Fig. 8.2. In the right plot of Fig. 8.2, the corresponding total widths of the intervals, i.e., the widths of the 90 % HDRs, are given. The strongest constraints are found for $\tilde{C}_{uW}$, with the width of its 90 % HDR being about 6×10^{-2}. As already observed in Chap. 6, the tight constraints on this coefficient originate from the measurements of W boson helicity fractions. Also $\tilde{C}_{uG}$ and $\tilde{C}_{uB}$ are well-constrained with the total widths of their 90 % HDRs being about 3×10^{-1} and 6×10^{-1}, due to the $t\bar{t}$ and $t\bar{t}\gamma$ cross section measurements, respectively. The coefficient $\tilde{C}_{\varphi q}^{(3)}$ is constrained on the order $\mathcal{O}(1)$ by the top-quark decay width measurement. The weakest constraints are obtained for $\tilde{C}_{\varphi q}^{(1)}$ and $\tilde{C}_{\varphi u}$. This is because only the $t\bar{t}Z$ cross section gives sensitivity to these two coefficients and the combined experimental and theory uncertainties of the corresponding measurement are rather large with a size of about 14 %.

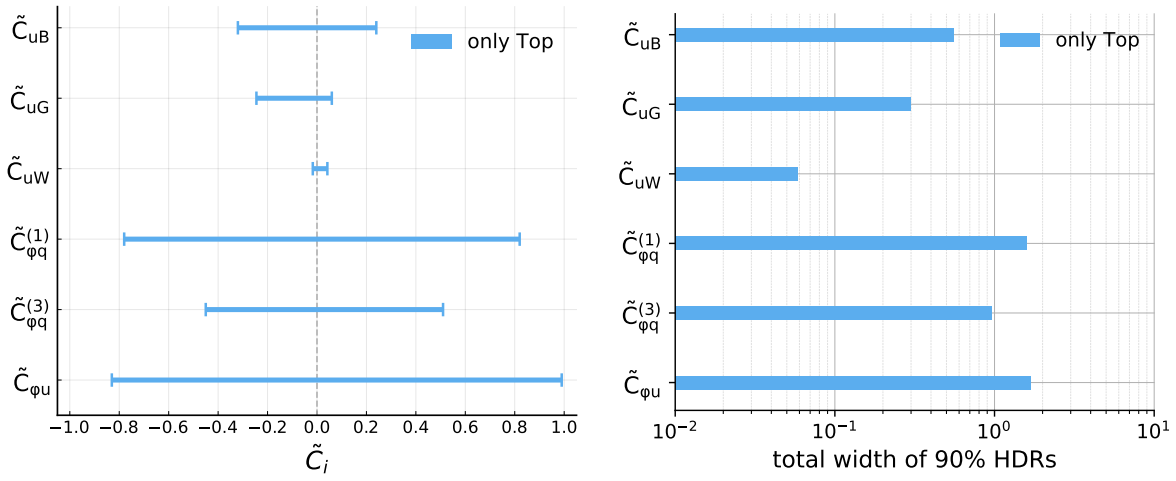

Figure 8.2: One-dimensional marginalized smallest intervals containing 90 % of the posterior probability (left) and their total widths (right) obtained in a fit of the Wilson coefficients in Eq. (8.4) to the top-quark data of Tab. 8.2.

8.3.2 Constraints from B and $Z \to b\bar{b}$ measurements

For the B physics observables, measurements from Belle, BaBar, LHCb, as well as averaged values from the HFLAV group, are considered here. The specific observables and corresponding references for the measurements and SM predictions are listed in Tab. 8.3.

Table 8.3: Considered observables of B physics processes and references for the corresponding measurements and SM predictions. For observables measured in q^2 bins, where q^2 denotes the square of the invariant dilepton mass, only one bin is included to minimize the effects of unknown bin-to-bin correlations.

Process	Observable	q^2 bin [GeV^2]	Experiment	Ref.	SM Ref.
$\bar{B} \to X_s\gamma$	$\mathrm{BR}_{E_\gamma > 1.6\ \mathrm{GeV}}$	-	HFLAV	[216]	[208]
$B^0 \to K^*\gamma$	BR	-	HFLAV	[216]	[200]
$B^+ \to K^{*+}\gamma$	BR	-	HFLAV	[216]	[200]
$\bar{B} \to X_s\ell^+\ell^-$	BR	$[1, 6]$	BaBar	[233]	[234]
	A_{FB}	$[1, 6]$	Belle	[235]	[234]
$B_s \to \mu^+\mu^-$	BR	-	LHCb	[236]	[200]
$B^0 \to K^*\mu^+\mu^-$	F_L, P_1, P_2, P_3, P_4', P_5', P_6', P_8'	$[1.1, 6]$	LHCb	[237]	[200]
$B^0 \to K^0\mu^+\mu^-$	$\mathrm{dBR}/\mathrm{d}q^2$	$[1, 6]$	LHCb	[238]	[200]
$B^+ \to K^+\mu^+\mu^-$	$\mathrm{dBR}/\mathrm{d}q^2$	$[1, 6]$	LHCb	[238]	[200]
$B^+ \to K^{+*}\mu^+\mu^-$	$\mathrm{dBR}/\mathrm{d}q^2$	$[1, 6]$	LHCb	[238]	[200]
$B_s \to \phi\mu^+\mu^-$	F_L, S_3, S_4, S_7	$[1, 6]$	LHCb	[239]	[200]
$\Lambda_b \to \Lambda\mu^+\mu^-$	$\mathrm{dBR}/\mathrm{d}q^2$	$[15, 20]$	LHCb	[240]	[200]
$B_s - \bar{B}_s$ mixing	ΔM_s	-	HFLAV	[216]	[226]

A fit of the eight SMEFT coefficients in Eq. (8.7) to the data of Tab. 8.3 is performed, again using a uniform prior distribution with $-1 \leq \tilde{C}_i \leq 1$ for all coefficients and the Metropolis–Hastings algorithm for sampling the posterior. The smallest 90 % intervals of the one-dimensional marginalized posterior distributions and their corresponding total widths are shown in Fig. 8.3 in light gray color.

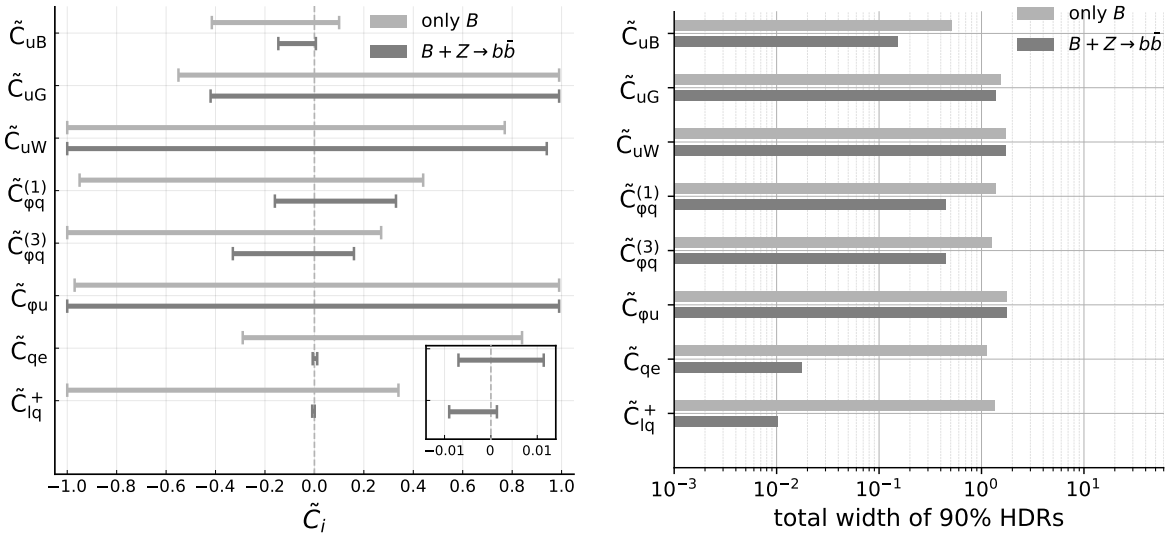

Figure 8.3: Marginalized smallest intervals containing 90 % of the posterior probability (left) and their total widths (right) obtained in fits of the SMEFT coefficients in Eq. (8.7) to only the B physics data of Tab. 8.3 (light gray) and to the combination of B and $Z \to b\bar{b}$ data in Tab. 8.4 (dark gray).

Only weak constraints on the eight Wilson coefficients are obtained. It is noticeable that in a fit using only the B physics data, $\tilde{C}_{uB}$ is the coefficient obtaining the tightest constraints, with the width of the corresponding 90 % HDR being of the order $\mathcal{O}(10^{-1})$. All other coefficients remain almost unconstrained in this fit, with the widths of the 90 % HDRs being of the order $\mathcal{O}(1)$. The main reason for the weak constraints in this fit to only the B physics data is that correlations between the coefficients are induced when matching WET and SMEFT. Considering the matching conditions in App. D.1, it can be noticed that there are more degrees of freedom in the SMEFT parametrization than in the WET parametrization. When performing the fit to the B data of Tab. 8.3 not in terms of the SMEFT coefficients, but in terms of WET coefficients at the scale μ_W, less correlated constraints are obtained. The results of such a fit to the WET coefficients are shown in Fig. D.3 of App. D. In this WET fit, it is observed that the obtained 90 % HDR of ΔC_9, i.e., the BSM part of $\bar{C}_9$, does not include the SM value. This deviation is a consequence of the B anomalies measured in the angular distributions of $B^0 \to K^* \mu^+ \mu^-$ decays and has been discussed in more detail in Ref. [241]. In the fit of the SMEFT coefficients shown in Fig. 8.3, however, this deviation is not observed since the constraints in SMEFT are diluted by the matching of WET and SMEFT.

As mentioned above, for improving the constraints on $\tilde{C}^{(1)}_{\varphi q}$ and $\tilde{C}^{(3)}_{\varphi q}$, also $Z \to b\bar{b}$ processes are considered. The specific observables included are listed in Tab. 8.4, together with corresponding references for the measurements and SM predictions.

Table 8.4: Included observables of $Z \to b\bar{b}$ processes and corresponding references for the measurements and SM predictions.

Process	Observables	Experiments	Refs.
$Z \to b\bar{b}$	$A_{\mathrm{FB}}^{0,b}$, R_b, $\sigma_{\mathrm{had.}}$	LEP/SLC	[101, 242]

A fit using the data of Tab. 8.4 can only constrain the linear combination $\tilde{C}_{\varphi q}^{+}$. The resulting posterior distribution of such a one-dimensional fit is shown in the left plot of Fig. 8.4. It is visible that the $Z \to b\bar{b}$ measurements yield tight constraints on $\tilde{C}_{\varphi q}^{+}$, with the width of its 90 % HDR being about 2×10^{-3}. Due to deviations from the SM in the measurement of the forward-backward asymmetry, the corresponding 90 % interval does not cover the SM, and a small deviation of about $2\,\sigma$ is obtained for the mode of $\tilde{C}_{\varphi q}^{+}$.

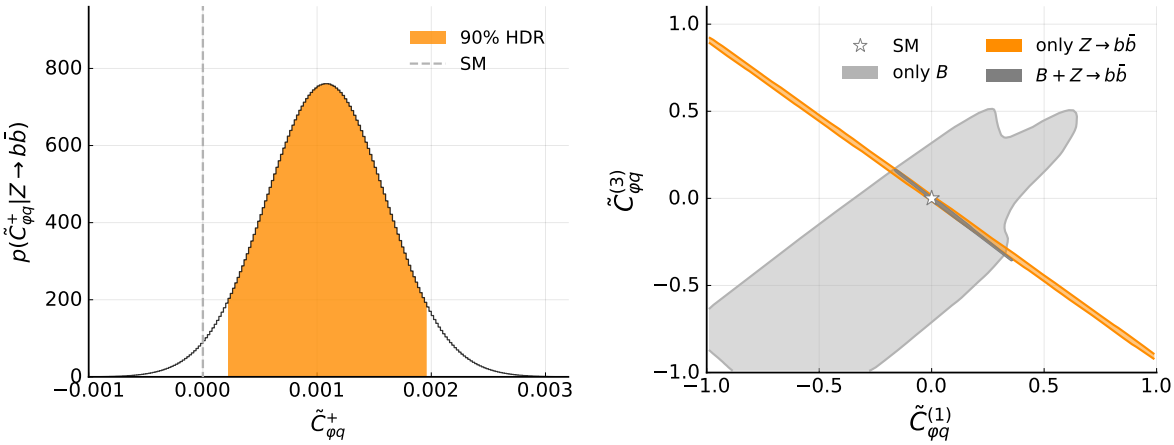

Figure 8.4: Posterior distribution of a one-dimensional fit of $\tilde{C}_{\varphi q}^{+}$ to the $Z \to b\bar{b}$ data of Tab. 8.4 (left) and two-dimensional marginalized 90 % HDRs (right) of $\tilde{C}_{\varphi q}^{(1)}$ vs. $\tilde{C}_{\varphi q}^{(3)}$ obtained in fits of the eight coefficients in Eq. (8.7) to only the B data (gray), to only the $Z \to b\bar{b}$ data (orange), and to their combination (dark gray).

When performing a fit of all eight Wilson coefficients in Eq. (8.7) to only the $Z \to b\bar{b}$ data of Tab. 8.4, the resulting one-dimensional marginal distributions yield no constraints and are flat for all coefficients. However, since the $Z \to b\bar{b}$ observables are sensitive to the linear combination $\tilde{C}_{\varphi q}^{+}$, a strong correlation in the two-dimensional marginal distribution of $\tilde{C}_{\varphi q}^{(1)}$ vs. $\tilde{C}_{\varphi q}^{(3)}$ is observed, and only a very narrow band is allowed in this projection of the parameter space. This is visualized in the right plot of Fig. 8.4, where the 90 % HDR of the corresponding two-dimensional marginal distribution of $\tilde{C}_{\varphi q}^{(1)}$ vs. $\tilde{C}_{\varphi q}^{(3)}$ is shown in the orange color. Performing a fit of the eight Wilson coefficients in Eq. (8.7) to the combination of the $Z \to b\bar{b}$ measurements of Tab. 8.4 and the B measurements of Tab. 8.3, the resulting constraints on $\tilde{C}_{\varphi q}^{(1)}$ and $\tilde{C}_{\varphi q}^{(3)}$ are improved significantly. In the right plot of Fig. 8.4, the two-dimensional 90 % HDRs of $\tilde{C}_{\varphi q}^{(1)}$ vs. $\tilde{C}_{\varphi q}^{(3)}$ are shown for fits of all eight coefficients to only the B physics data (light gray), to only the $Z \to b\bar{b}$ data (orange), and to their combination (dark gray). It is visible that the individual data sets have orthogonal sensitivities to the coefficients, which

yields the tight constraints in their combination. The improved constraints on $\tilde{C}_{\varphi q}^{(1)}$ and $\tilde{C}_{\varphi q}^{(3)}$ obtained in a combined fit to B and $Z \to b\bar{b}$ data are also observed in the one-dimensional constraints in Fig. 8.3, shown in the dark gray color. There, the constraints of $\tilde{C}_{\varphi q}^{(1)}$ and $\tilde{C}_{\varphi q}^{(3)}$ improve by more than a factor of two, compared to the fit using only the B physics data. Furthermore, in Fig. 8.3, an improvement by about two orders of magnitude is striking for the constraints on $\tilde{C}_{qe}$ and $\tilde{C}_{lq}^{+}$ in the combined fit. As the $Z \to b\bar{b}$ data is not directly sensitive to these two coefficients, this improvement is initially unexpected. However, it can be explained when taking into account that the constraints shown in Fig. 8.3 are only marginalized views of the total eight-dimensional parameter space. Unfortunately, visualizing the complete multidimensional posterior distributions is not possible. However, considering three-dimensional visualizations of the posterior distributions already allows explaining the tight constraints on $\tilde{C}_{qe}$ and $\tilde{C}_{lq}^{+}$ observed in the combination. In Fig. 8.5, three-dimensional posterior distributions are shown for the individual and the combined fits of the eight Wilson coefficients. The marginal views of $\tilde{C}_{\varphi q}^{(1)}$ vs. $\tilde{C}_{\varphi q}^{(3)}$ vs. $\tilde{C}_{qe}$ (left) and $\tilde{C}_{\varphi q}^{(1)}$ vs. $\tilde{C}_{\varphi q}^{(3)}$ vs. $\tilde{C}_{lq}^{+}$ (right) are depicted. In both cases, it can be observed how the constraints from the B physics data intersect the plane-shaped constraints from the $Z \to b\bar{b}$ measurements. In the one- and two-dimensional marginal distributions, these orthogonalities in the three-dimensional parameter space are not visible because the spatial information is lost due to the marginalization. When considering these three-dimensional visualizations, however, it becomes apparent how the orthogonality of the constraints in the multidimensional parameter space leads to the tight constraints on the coefficients $\tilde{C}_{qe}$ and $\tilde{C}_{lq}^{+}$ in the combination of B and $Z \to b\bar{b}$ data. Thus, even if certain observables, such as the $Z \to b\bar{b}$ observables in this example, do not provide tight constraints on the considered coefficients, they can still exclude a large fraction of the allowed parameter space. This also illustrates how combining observables sensitive to specific coefficients can significantly improve the constraints on other coefficients, as the complementary sensitivities in the multidimensional parameter space have to be considered.

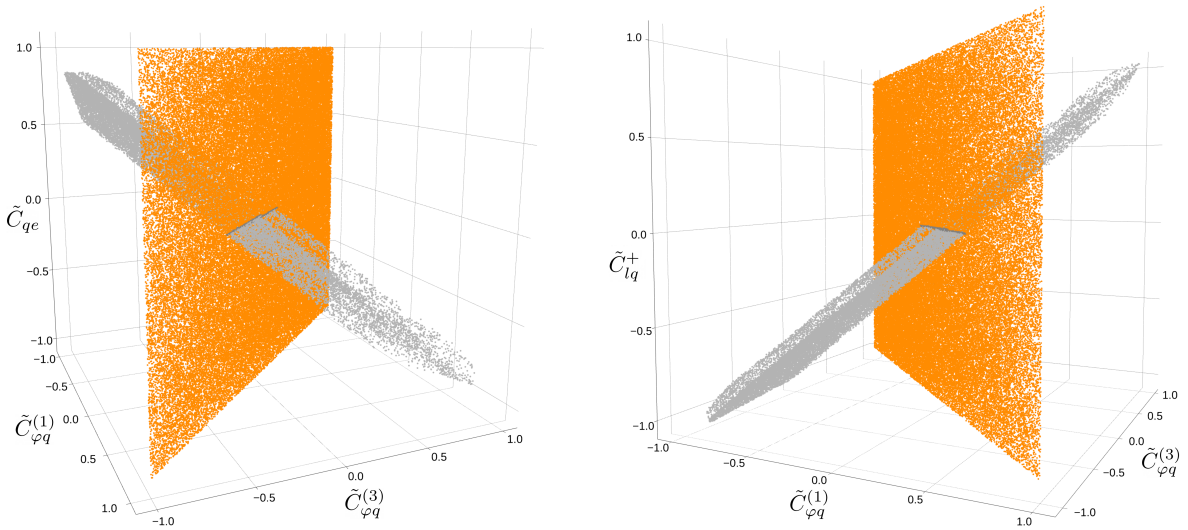

Figure 8.5: Three-dimensional marginalized posterior distributions for $\tilde{C}_{\varphi q}^{(1)}$ vs. $\tilde{C}_{\varphi q}^{(3)}$ vs. $\tilde{C}_{qe}$ (left) and for $\tilde{C}_{\varphi q}^{(1)}$ vs. $\tilde{C}_{\varphi q}^{(3)}$ vs. $\tilde{C}_{lq}^{+}$ (right) obtained in fits of the eight coefficients of Eq. (8.7) to only the B data (light gray), to only the $Z \to b\bar{b}$ data (orange), and to their combination (dark gray).

8.3.3 Constraints from a combined fit

A combined fit of the eight coefficients in Eq. (8.7) to the top-quark, B, and $Z \to b\bar{b}$ measurements is performed. Again, uniform prior distributions in the range $-1 \leq \tilde{C}_i \leq 1$ are applied for all coefficients. The ellipsoidal nested sampling algorithm of BAT.jl is used for exploring the multimodal posterior distribution. In Sec. 8.5, it is discussed in more detail why nested sampling algorithms are required for exploring these posterior distributions. The resulting smallest 90 % intervals of the combined fit are given in the left plot of Fig. 8.6 in dark blue color. Also indicated are the constraints from a fit to only the top-quark data (light blue) and from the fit to only the B and $Z \to b\bar{b}$ data (dark gray), as already shown in Fig. 8.3. In the right plot of Fig. 8.6, the corresponding total widths of the 90 % HDRs are given.

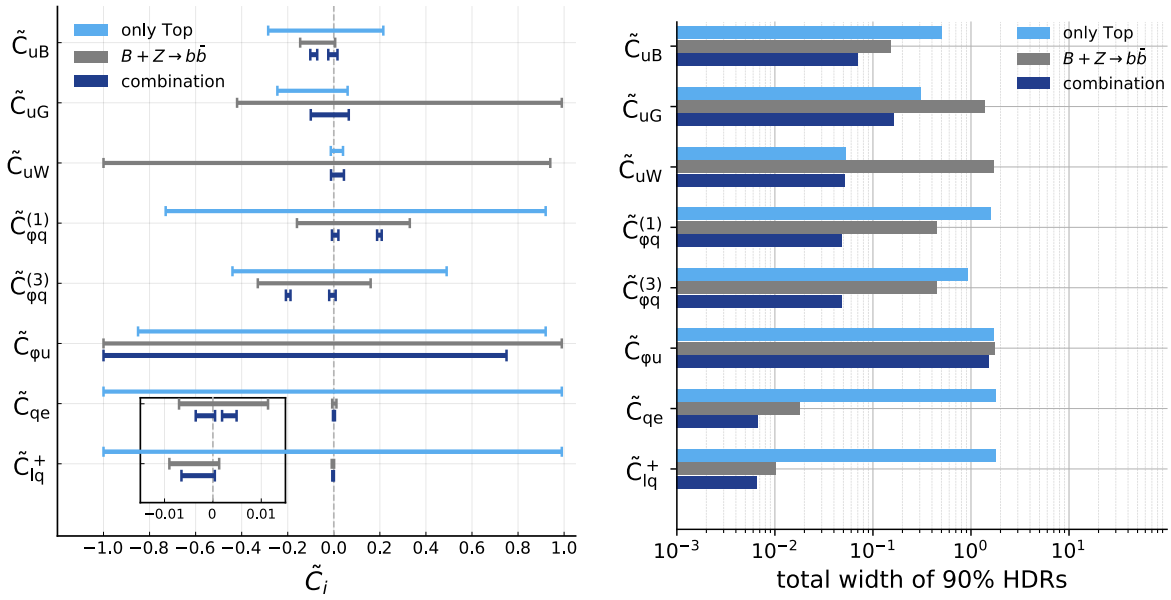

Figure 8.6: Marginalized smallest intervals containing 90 % of the posterior probability (left) and their total widths (right) obtained in fits of the Wilson coefficients in Eq. (8.7) to only the top-quark data (light blue), to only the combined B and $Z \to b\bar{b}$ data (gray), and to their combination (dark blue).

In the combined fit, most of the Wilson coefficients are constrained tightly. Only $\tilde{C}_{\varphi u}$ remains essentially unconstrained, as only the $t\bar{t}Z$ cross section is sensitive to this coefficient, and even the combination cannot improve the weak constraints on this coefficient, which were already observed in the fit to only the top-quark data in Fig. 8.2. For the coefficients $\tilde{C}_{uB}$, $\tilde{C}_{uG}$, $\tilde{C}_{uW}$, $\tilde{C}^{(1)}_{\varphi q}$, and $\tilde{C}^{(3)}_{\varphi q}$, tight constraints are obtained with the width of the corresponding HDRs being about 5×10^{-2}. These constraints are up to an order of magnitude smaller than those obtained in the fits to the individual data sets. For $\tilde{C}_{qe}$ and for the linear combination $\tilde{C}^{+}_{lq}$, the constraints in the combined fit are at a level of 6×10^{-3}, which is a consequence of the already tight constraints obtained in the combination of B and $Z \to b\bar{b}$ data. The considered top-quark data cannot constrain these coefficients since none of the top-quark observables is sensitive to them, as discussed in Sec. 8.2. For $\tilde{C}^{(1)}_{\varphi q}$ and $\tilde{C}^{(3)}_{\varphi q}$, the improvements of the constraints in the combination are the largest. For both coefficients, the 90 % HDRs consist of two clearly

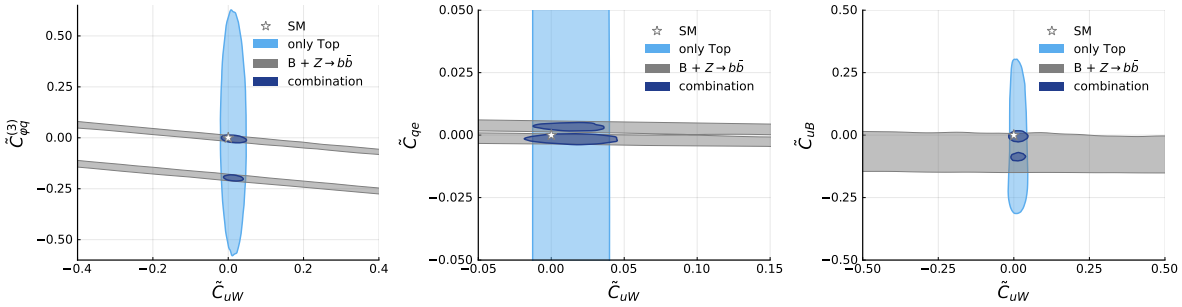

Figure 8.7: Two-dimensional marginalized posterior distributions obtained in fits of the Wilson coefficients in Eq. (8.7) to only the top-quark data (light blue), to only the combination of B and $Z \to b\bar{b}$ data (gray) and to the total combination (dark blue). Shown are the 90 % HDRs of the posteriors for $\tilde{C}_{uW}$ vs. $\tilde{C}^{(3)}_{\varphi q}$, $\tilde{C}_{uW}$ vs. $\tilde{C}_{qe}$, and $\tilde{C}_{uW}$ vs. $\tilde{C}_{uB}$. The SM value is indicated by the star.

separated regions, with one of them being in agreement with the SM. Also for $\tilde{C}_{uB}$ and $\tilde{C}_{qe}$, such separated HDRs are obtained. Again, this cannot be explained by considering only the one-dimensional 90 % HDRs shown in Fig. 8.3. However, the origin of the divided HDRs in the combination can be understood when considering the corresponding two-dimensional marginal distributions. In Fig. 8.7, the two-dimensional marginal distributions are given for $\tilde{C}_{uW}$ vs. $\tilde{C}^{(3)}_{\varphi q}$, $\tilde{C}_{uW}$ vs. $\tilde{C}_{qe}$, and $\tilde{C}_{uW}$ vs. $\tilde{C}_{uB}$. The 90 % HDRs obtained in the fits to only the top-quark data (light blue), to the combination of B and $Z \to b\bar{b}$ data (gray), and to the combination of all measurements (dark blue) are shown. In all three plots, it is observed that the constraints from the top-quark and from the combined B and $Z \to b\bar{b}$ data are orthogonal in the considered projections. As discussed before, this leads to a significant reduction of the allowed parameter space and to the observed improvement in the combination of all observables. In particular, in the left plot, it can be noticed that the constraints from the B and $Z \to b\bar{b}$ measurements are multimodal in $\tilde{C}^{(3)}_{\varphi q}$, which is not visible in the one-dimensional 90 % HDRs of Fig. 8.6, as this spatial information is again lost due to marginalization. In the combination, the constraints from the top-quark data intersect with the flat constraints from the combined B and $Z \to b\bar{b}$ data, and the orthogonality of the constraints leads to the two solutions in $\tilde{C}^{(3)}_{\varphi q}$ being also visible in the one-dimensional HDRs. In the other two plots, a similar effect is observed, but the two solutions in the B and $Z \to b\bar{b}$ constraints are much closer together. While in the middle plot of $\tilde{C}_{uW}$ vs. $\tilde{C}_{qe}$ a separation of the constraints can still be seen, in particular for the right plot of $\tilde{C}_{uW}$ vs. $\tilde{C}_{uB}$, the bimodality can only be resolved in the combined fit.

As such ambiguous solutions with separated HDRs cannot be resolved by current data alone, in the following, projections for future measurements are investigated for their potential to further improve the constraints on the SMEFT coefficients and to exclude these second solutions.

8.4 Prospects for SMEFT fits with future measurements

In the coming years, much more data is going to be analyzed by the LHC experiments [243], and new precision measurements of B observables are expected to be performed by the Belle II

experiment [244]. As already indicated in Chap. 6, improved precision of the measurements will lead to tighter constraints on the SMEFT coefficients and could allow detecting deviations from the SM. Here, two future scenarios are considered to investigate the possible effects of prospective measurements on the constraints of SMEFT Wilson coefficients, in particular with respect to the synergies of top-quark and B physics observables. In a near-future scenario, expected results from the ATLAS and CMS experiments at the HL-LHC and from the Belle II experiment are considered. While the HL-LHC experiments are foreseen to start taking data by the end of this decade, the Belle II experiment has already started data taking in 2019. Both facilities will provide measurements of the observables considered above with improved precision. Moreover, measurements of processes that have not been observed before are anticipated. Specifically, measurements of $b \to s\nu\bar{\nu}$ transitions are expected to be possible at Belle II and are explicitly taken into account in the following studies. Currently, only upper limits on such processes have been derived [245]. In a second step, a far-future projection is studied. In this scenario, the opportunities for improving SMEFT fits with measurements performed at a future lepton collider are investigated. So far, there has not been a decision on a specific lepton collider, and different concepts, such as the *Future Circular Collider* (FCC)-ee [246], the *International Linear Collider* (ILC) [247], and the *Compact Linear Collider* (CLIC) [248], are still under discussion. In the following studies, the CLIC proposal is adopted for the far-future scenario, and possible improvements of SMEFT fits by including measurements of top-quark observables expected at such a lepton collider are investigated.

The specific assumptions made for the future projections of both scenarios are discussed in the following paragraphs, and their effects on the constraints of SMEFT coefficients are explored.

8.4.1 Near-future scenario

In the near-future scenario, precise measurements of top-quark and B physics observables are considered to be performed by the ATLAS and CMS experiments at the HL-LHC and by the Belle II experiment at the SuperKEKB collider. The approach employed here for estimating future projections is similar to that used in Chap. 6. The nominal values of the future measurements are assumed to be those of the current measurements given in Sec. 8.3. For observables that have not yet been measured, the corresponding SM predictions are taken as the nominal values. The simplifying assumption of SM theory uncertainties decreasing by a factor of two compared to the current SM predictions is made to account for expected improvements in MC simulations and higher-order theory calculations. Estimates of the values of experimental uncertainties, i.e., statistical and systematic uncertainties, are based on studies of the expected precision by the ATLAS, CMS, and Belle II collaborations in Refs. [243, 244, 249–251]. The statistical uncertainties of the measurements are estimated by scaling current uncertainties according to the expected integrated luminosity. For the HL-LHC experiments, an integrated luminosity of $3000\,\text{fb}^{-1}$ is assumed for a center-of-mass energy of 14 TeV [243]. For Belle II, an integrated luminosity of $50\,\text{ab}^{-1}$ is anticipated [244]. The specific processes and observables considered in the near-future scenario, as well as the corresponding references used for estimating the expected precision, are given in Tab. 8.5.

Table 8.5: Observables expected to be measured at the HL-LHC experiments and Belle II in the near-future scenario. The references considered for estimating the expected precision and the SM predictions are given.

Process	Observable	q^2 bin [GeV^2]	Experiment	Ref.	SM Ref.
$t\bar{t}\gamma$	$\sigma^{\text{fid}}(t\bar{t}\gamma, 1\ell)$, $\sigma^{\text{fid}}(t\bar{t}\gamma, 2\ell)$	-	ATLAS	[243, 249]	[206, 252]
$t\bar{t}Z$	$\sigma^{\text{inc}}(t\bar{t}Z)$	-	CMS	[243, 250]	[229–231]
$t\bar{t}$	$\sigma^{\text{inc}}(t\bar{t})$	-	CMS	[243, 251]	[232]
	F_0 , F_L	-	-	-	[177]
	Γ_t	-	-	-	[184]
$\bar{B} \to X_s\gamma$	$\text{BR}_{E_\gamma > 1.6\text{ GeV}}$	-	Belle II	[244]	[208]
$B^0 \to K^*\gamma$	BR	-	Belle II	[244]	[200]
$B^+ \to K^{+*}\gamma$	BR	-	Belle II	[244]	[200]
$\bar{B} \to X_s\ell^+\ell^-$	BR, A_{FB}	$[3.5, 6]$	Belle II	[244]	[234]
$B^0 \to K^*\mu^+\mu^-$	F_L , P_1 , P_2 , P_3 , P_4' , P_5' , P_6' , P_8'	$[1.1, 2.5]$, $[2.5, 4]$, $[4, 6]$	Belle II	[244]	[200]
$B^0 \to K^{(*)}\nu\bar{\nu}$	BR	-	Belle II	[244]	[200]

For the fiducial $t\bar{t}\gamma$ cross sections, estimates of the anticipated uncertainties are given in Refs. [243, 249]. For both the single-lepton and the dilepton channel, the largest estimates of the experimental uncertainties are assumed here. For the $t\bar{t}Z$ cross section, the projections of Refs. [243, 250] are considered, and the statistical uncertainties are scaled to the corresponding luminosity. For the total $t\bar{t}$ cross section, the assumed precision is based on the studies of differential cross sections in Refs. [243, 251]. For both the $t\bar{t}Z$ and the $t\bar{t}$ cross sections, the systematic uncertainties are assumed to decrease by a factor of two compared to the current values. As expectations for the top-quark decay observables are not discussed in Ref. [243], the uncertainties are estimated similarly to those of the $t\bar{t}$ cross section. For future projections of $b \to s$ transition measurements by the Belle II experiment, the estimates in Ref. [244] are considered. For $\text{BR}(\bar{B} \to X_s\gamma)_{E_\gamma > 1.6\,\text{GeV}}$, it is assumed that the given uncertainties for $E_\gamma > 1.9\,\text{GeV}$ can be applied. In the case of the [1.1, 2.5] q^2-bin for $B^0 \to K^*\mu^+\mu^-$, the Belle II projection of the [1.0, 2.5] bin is considered for the expected experimental uncertainties. In particular, it is assumed that Belle II yields first observations of $b \to s\nu\bar{\nu}$ transitions by measuring the branching ratio of $B^0 \to K^{(*)}\nu\bar{\nu}$. As indicated in Tab. 8.1, this observable is sensitive to the linear combinations $\tilde{C}^+_{\varphi q}$ and $\tilde{C}^-_{lq}$. The current measurements of Sec. 8.3 are sensitive to the linear combination $\tilde{C}^+_{lq}$, i.e., the linear combination with the opposite sign as $\tilde{C}^-_{lq}$. Therefore, by including $b \to s\nu\bar{\nu}$ transitions in the SMEFT fits, it is expected that $\tilde{C}^{(1)}_{lq}$ and $\tilde{C}^{(3)}_{lq}$ can be constrained individually. The dependences of the top-quark and B physics observables on the considered Wilson coefficients are computed as described in Sec. 8.2. For the new observable $\text{BR}(B^0 \to K^{(*)}\nu\bar{\nu})$, Ref. [253] is followed for the computation of BSM contributions, as detailed in Ref. [5].

A combined fit to the current measurements of Sec. 8.3 and to the near-future projections of Tab. 8.5 is performed with EFTfitter.jl. Again, uniform prior distributions with maximum

ranges of $|\tilde{C}_i| \leq 1$ are chosen, and the ellipsoidal nested sampling algorithm is used for exploring the posterior distribution. Nine SMEFT coefficients are constrained, i.e., those of Eq. (8.7) but including $\tilde{C}_{lq}^{(1)}$ and $\tilde{C}_{lq}^{(3)}$ individually instead of using their linear combination $\tilde{C}_{lq}^{+}$. The constraints on the coefficients obtained in this fit are shown in Fig. 8.8 in red color. Again, the marginalized smallest intervals containing 90 % posterior probability (left) are shown together with the corresponding total widths of these intervals (right). Also shown (in blue color) are the constraints obtained in a fit of the nine coefficients to only the current measurements.

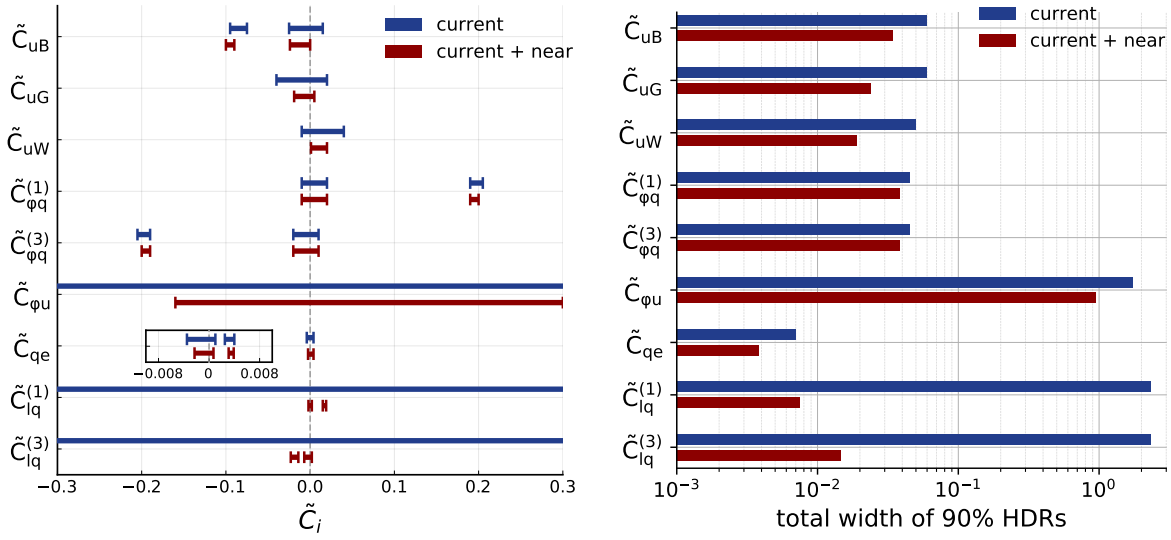

Figure 8.8: Marginalized smallest intervals containing 90 % of the posterior probability (left) and their total widths (right) obtained in fits of nine Wilson coefficients when using only the current measurements (blue) and when using the combination of current and near-future measurements (red).

It is visible that when using the combination of current and near-future measurements, almost all Wilson coefficients are constrained tightly. As in the previous fits, only for $\tilde{C}_{\varphi u}$ weak constraints of the order $\mathcal{O}(1)$ are obtained because still only the $t\bar{t}Z$ cross section gives sensitivity to this coefficient. It has to be noted that in the left plot of Fig. 8.8, only one of the two 90 % HDRs of $\tilde{C}_{\varphi u}$ is visible for the combined current and near-future fit, as the other part is in the range $-1 \leq \tilde{C}_{\varphi u} \leq -0.6$ and is excluded for better visibility of the constraints on the other coefficients. The one-dimensional marginalized posterior distribution of this coefficient is shown in Fig. 8.10, where the second branch of the 90 % HDR is also visible. In the combined fit, the total widths of the constraints on $\tilde{C}_{uB}$, $\tilde{C}_{uG}$, $\tilde{C}_{uW}$, $\tilde{C}_{\varphi q}^{(1)}$, and $\tilde{C}_{\varphi q}^{(3)}$ are all in the range of $(2-4)\times 10^{-2}$. Compared to the fit using only the current measurements, the constraints on $\tilde{C}_{uB}$, $\tilde{C}_{uG}$, $\tilde{C}_{uW}$ are tightened by about a factor of two, while the constraints on $\tilde{C}_{\varphi q}^{(1)}$ and $\tilde{C}_{\varphi q}^{(3)}$ remain almost unchanged. As discussed in Sec. 8.2, using only current data, the coefficients $\tilde{C}_{lq}^{(1)}$ and $\tilde{C}_{lq}^{(3)}$ cannot be constrained independently, but only their linear combination $\tilde{C}_{lq}^{+}$, to which measurements of $b \to s\ell^+\ell^-$ observables provide sensitivity (as shown in Fig. 8.6). As visible in Fig. 8.8, when combined with the anticipated Belle II measurement of the $b \to s\nu\bar{\nu}$ observable, sensitive to the opposite linear combination $\tilde{C}_{lq}^{-}$, the coefficients $\tilde{C}_{lq}^{(1)}$ and $\tilde{C}_{lq}^{(3)}$ can

be probed independently and tight constraints in the range of $(7-11)\times10^{-3}$ for the widths of the 90 % HDRs are obtained. The origin of these tight constraints on the four-fermion operators in the combination of current and near-future measurements is visualized in Fig. 8.9. In the left plot, the 90 % HDRs obtained in fits to only the current data (blue), to only the near-future data (orange), and to their combination (red) are shown.

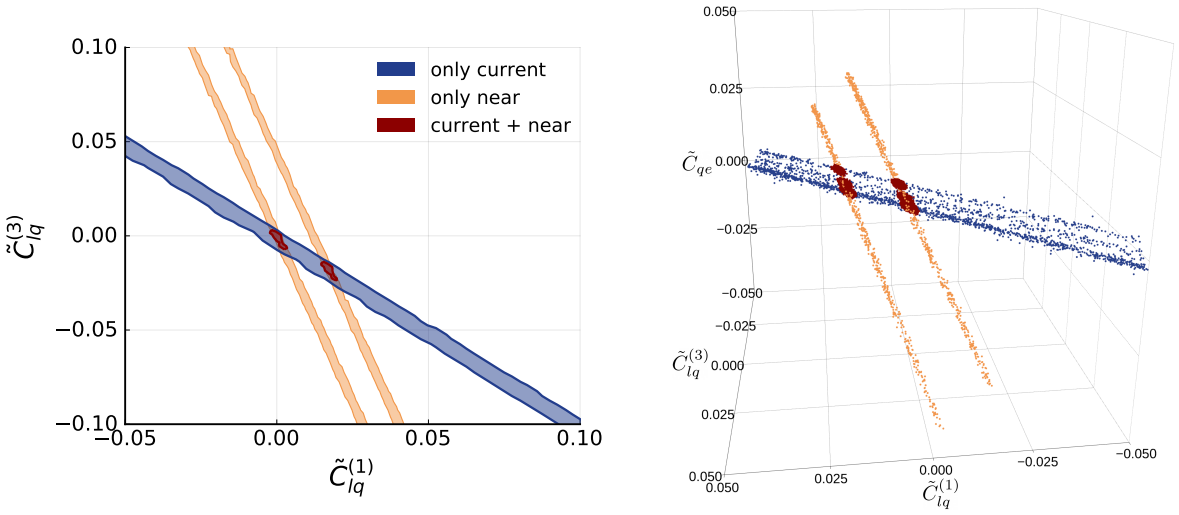

Figure 8.9: Marginalized two-dimensional 90 % HDRs of $\tilde{C}_{lq}^{(1)}$ vs. $\tilde{C}_{lq}^{(3)}$ (left) and marginalized three-dimensional posterior distribution of $\tilde{C}_{lq}^{(1)}$ vs. $\tilde{C}_{lq}^{(3)}$ vs. $\tilde{C}_{qe}$ (right) obtained in fits of nine Wilson coefficients to only the current measurements (blue), to only the near-future measurements (orange), and to their combination (red).

It is visible that due to the inclusion of the $b \to s\nu\bar{\nu}$ observable, the current and the near-future measurements have complementary sensitivities to the Wilson coefficients $\tilde{C}_{lq}^{(1)}$ and $\tilde{C}_{lq}^{(3)}$. Their combination, therefore, yields the observed tight constraints. In the right plot of Fig. 8.9 this is also shown in terms of the three-dimensional marginal distribution of $\tilde{C}_{lq}^{(1)}$ vs. $\tilde{C}_{lq}^{(3)}$ vs. $\tilde{C}_{qe}$. There, it is visible that the improved sensitivity to $\tilde{C}_{lq}^{(1)}$ and $\tilde{C}_{lq}^{(3)}$ also significantly narrows the allowed region in the multidimensional parameter space and, thus, improves the constraints on $\tilde{C}_{qe}$. This is a similar effect as the one observed in Sec. 8.3.2 in the combination of B and $Z \to b\bar{b}$ measurements, but not quite as strong.

As can be seen in Fig. 8.8, for most coefficients, the 90 % HDRs are split into two parts. For $\tilde{C}_{uB}$, $\tilde{C}_{\varphi q}^{(1)}$, $\tilde{C}_{\varphi q}^{(3)}$, and $\tilde{C}_{qe}$, this was already observed in the fit to current data in Fig. 8.6. For the coefficients $\tilde{C}_{\varphi u}$, $\tilde{C}_{qe}$, $\tilde{C}_{lq}^{(1)}$, and $\tilde{C}_{lq}^{(3)}$, the one-dimensional marginalized distributions obtained in the combined current and near-future fit are shown in Fig. 8.10 and reveal the multimodal shape of the posterior distribution. For all coefficients, the parts of the 90 % HDRs that are closer to the SM value cover the largest proportion of the posterior probability, and thus all constraints are in good agreement with the SM.

In the next paragraph, it is investigated if including far-future measurements from a CLIC-like lepton collider can resolve these ambiguous solutions.

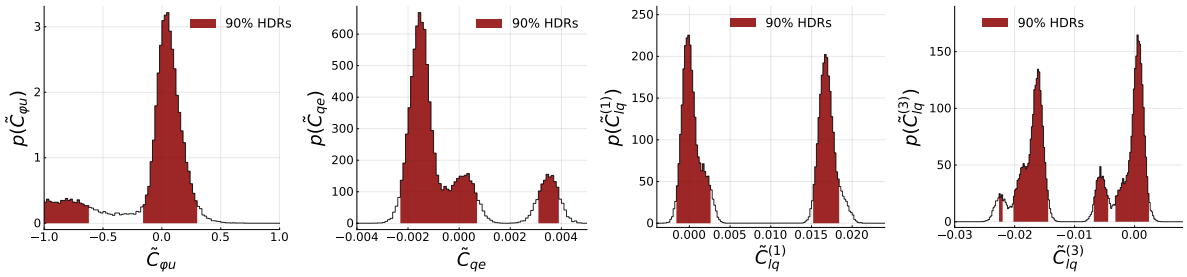

Figure 8.10: One-dimensional marginalized posterior distributions of $\tilde{C}_{\varphi u}$, $\tilde{C}_{qe}$, $\tilde{C}_{lq}^{(1)}$, and $\tilde{C}_{lq}^{(3)}$ obtained in the combined fit to current and near-future measurements. The smallest intervals containing 90 % of the posterior probability are highlighted.

8.4.2 Far-future scenario

In the far-future scenario, measurements of top-quark observables performed at the CLIC e^+e^-–collider are anticipated. Specifically, the total cross section of $t\bar{t}$ production $\sigma(t\bar{t})$ and the forward-backward asymmetry A_{FB} are considered here. As indicated in Tab. 8.1, such observables of $e^+e^- \to t\bar{t}$ processes are sensitive to the coefficients $\tilde{C}_{uB}$, $\tilde{C}_{uW}$, $\tilde{C}_{uG}$, $\tilde{C}_{\varphi u}$, $\tilde{C}_{qe}$, and to the linear combinations $\tilde{C}_{\varphi q}^{-}$ and $\tilde{C}_{lq}^{-}$. In addition, they also provide sensitivity to the previously unconstrained four-fermion coefficients $\tilde{C}_{eu}$ and $\tilde{C}_{lu}$.[1] In its current design proposal, different operation modes of CLIC are foreseen. On the one hand, three different center-of-mass energies are possible: 380 GeV, 1.4 TeV, and 3.0 TeV. On the other hand, two different beam polarizations are proposed, namely longitudinal polarizations of ±80 % for the electron beam and no polarization of the positron beam [248]. The observables and operation modes considered in the far-future scenario are given in Tab. 8.6.

Table 8.6: Observables and operation modes considered for $t\bar{t}$ production at CLIC in the far-future scenario.

Observable	$\sqrt{s}$	Polarization (e^-, e^+)	Ref. experiment	SM Ref.
$\sigma_{t\bar{t}}$, A_{FB}	380 GeV, 1.4 TeV, 3 TeV	$(\pm 80\,\%, 0)$	[248]	[255]

First studies on top-quark physics at CLIC have been presented in Ref. [248] and are followed here to obtain estimates for the expected precision of the measurements. For the nominal values, current SM predictions are assumed, including NLO QCD corrections [255]. The dependences of the cross section and the forward-backward asymmetry on the SMEFT coefficients are again determined using MC simulations with MadGraph5_aMC@NLO and subsequent interpolations, as described in Sec. 8.2.

When performing a fit with EFTfitter.jl to the expected CLIC measurements of Tab. 8.6 only, all of the eleven SMEFT coefficients in Fig. 8.1 are constrained. The marginalized 90 % HDRs and their total widths obtained in this fit are shown in Fig. 8.11 in the light brown color.

[1] During the finalization of these studies, a preprint by the CMS collaboration was published in which the coefficients $\tilde{C}_{eu}$ and $\tilde{C}_{lu}$ are constrained by measurements of top-quark production in association with leptons [254]. These two coefficients are otherwise unconstrained by current data.

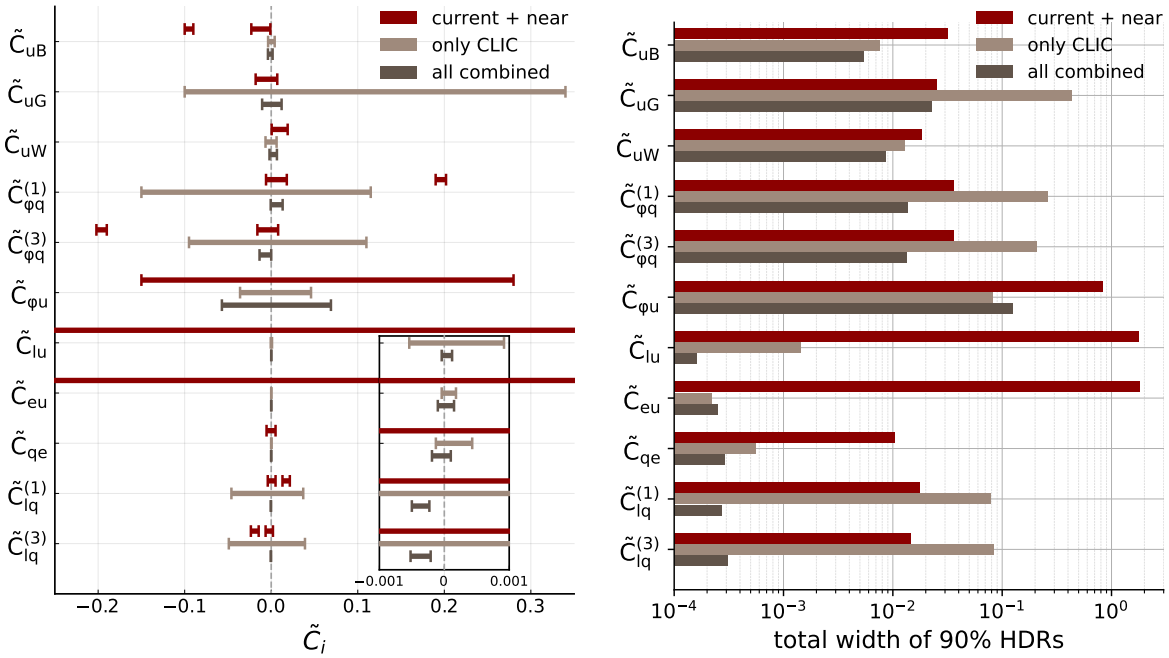

Figure 8.11: Marginalized smallest intervals containing 90 % of the posterior probability (left) and their total widths (right) obtained in fits of the eleven Wilson coefficients in Fig. 8.1 to only the combined current and near-future data (red), to only the CLIC data (light brown) and to the total combination (dark brown).

The tightest constraints are found for the four-fermion coefficients $\tilde{C}_{eu}$ and $\tilde{C}_{qe}$, with the widths of their 90 % HDRs being about $(3-6)\times10^{-4}$. Strong constraints on the order of $\mathcal{O}(10^{-3})$ are also obtained for $\tilde{C}_{lu}$. The weakest constraints in the fit to only the CLIC data are derived for $\tilde{C}_{uG}$. As indicated in Tab. 8.1, the sensitivity of the $e^+e^- \to t\bar{t}$ observables to this coefficient arises only through running in SMEFT and is therefore suppressed. Also for the coefficients $\tilde{C}^{(1)}_{\varphi q}$, $\tilde{C}^{(3)}_{\varphi q}$, $\tilde{C}^{(1)}_{lq}$, and $\tilde{C}^{(3)}_{lq}$, the constraints obtained in the fit to only the CLIC data are weaker than those obtained in a combined fit of all eleven coefficients to the current and near-future measurements of Sec. 8.4.1, as shown in Fig. 8.11 in the red color. This is due to the considered CLIC observables being only sensitive to the linear combinations $\tilde{C}^{-}_{\varphi q}$ and $\tilde{C}^{-}_{lq}$, as shown in Tab. 8.1. That these four coefficients can be constrained independently at all using only the CLIC data is an effect of the RGE running in which these coefficients evolve differently.[2] The constraints on the remaining coefficients, i.e., $\tilde{C}_{uB}$, $\tilde{C}_{uW}$, and $\tilde{C}_{\varphi u}$, are stronger than those obtained using the combined current and near-future measurements. In particular, the $e^+e^- \to t\bar{t}$ processes considered in the far-future scenario provide additional sensitivity to $\tilde{C}_{\varphi u}$ and thus allow improving the constraints by almost one order of magnitude. Also shown in Fig. 8.11 (in the dark brown color) are the constraints on the coefficients obtained in a fit to the combination of current, near-future, and far-future measurements. It is noticeable that

[2] It should be noted that the constraints on $\tilde{C}^{(1)}_{\varphi q}$ and $\tilde{C}^{(3)}_{\varphi q}$ shown in Fig. 8.11 are stronger than those shown in the corresponding plot of Ref. [5]. This is because in this thesis, the constraints on the SMEFT coefficients are given at the scale 172.5 GeV instead of at the scale 1 TeV as in Ref. [5]. Therefore, here, the effects of SMEFT running are more pronounced.

in this total combination, all coefficients can be constrained very tightly. The coefficients of the four-fermion operators are constrained the strongest, with the total widths of their 90 % HDRs being about $(2-3)\times10^{-4}$. In particular, for the coefficients $\tilde{C}_{lq}^{(1)}$ and $\tilde{C}_{lq}^{(3)}$, a decrease in the size of the 90 % HDRs by more than an order of magnitude is observed when compared to the fits using only the CLIC or only the combined current and near-future data sets. This large improvement is again a consequence of the complementary sensitivities of the considered observables. This can be observed particularly well in the left plot of Fig. 8.12, where the 90 % HDRs of the two-dimensional marginalized posterior distributions are shown for the individual and combined fits. Such orthogonal constraints obtained from the different data sets are also observed for further coefficients. In the right plot of Fig. 8.12 this is demonstrated for the two-dimensional marginal distribution of $\tilde{C}_{\varphi q}^{(1)}$ vs. $\tilde{C}_{lu}$.

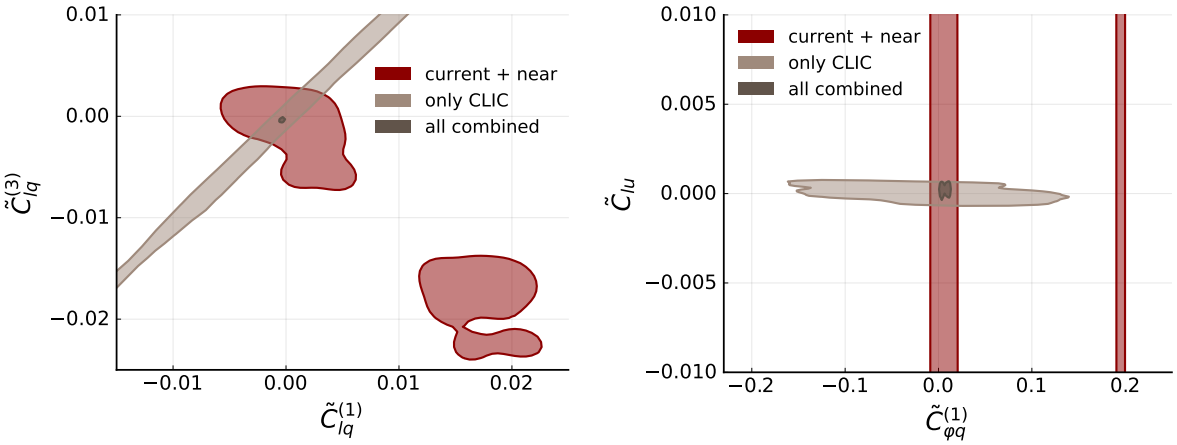

Figure 8.12: Two-dimensional 90 % HDRs of $\tilde{C}_{lq}^{(1)}$ vs. $\tilde{C}_{lq}^{(3)}$ (left) and $\tilde{C}_{\varphi q}^{(1)}$ vs. $\tilde{C}_{lu}$ (right) obtained in fits to the combined current and near-future measurements (red), to only the CLIC measurements (light brown), and to the total combination of all measurements (dark brown).

As noticeable in Fig. 8.11, for the coefficients $\tilde{C}_{uB}$, $\tilde{C}_{\varphi q}^{(1)}$, $\tilde{C}_{\varphi q}^{(3)}$, $\tilde{C}_{lq}^{(1)}$, and $\tilde{C}_{lq}^{(3)}$, the combination with the CLIC measurements allows removing the second solutions observed in the fits to the combined current and near-future data sets. This is again a result of the orthogonality of the constraints from the CLIC data and from the combined current and near-future data. In the plots of Fig. 8.12, it is visible how this complementarity of the constraints allows excluding the second solutions for the coefficients $\tilde{C}_{\varphi q}^{(1)}$, $\tilde{C}_{lq}^{(1)}$, and $\tilde{C}_{lq}^{(3)}$. Again, the observed synergies in the sensitivities of the different data sets also affect the multidimensional parameter space and allow tightening the constraints on further coefficients. Two examples of three-dimensional marginalized posterior distributions where this is visible are shown in Fig. 8.13 for $\tilde{C}_{\varphi q}^{(1)}$ vs. $\tilde{C}_{lu}$ vs. $\tilde{C}_{qe}$ (left) and $\tilde{C}_{lq}^{(1)}$ vs. $\tilde{C}_{qe}$ vs. $\tilde{C}_{lq}^{(3)}$ (right).

It should be noted that in Fig. 8.11, deviations from the SM are observed for $\tilde{C}_{lq}^{(1)}$ and $\tilde{C}_{lq}^{(3)}$ in the combined fit, with the corresponding modes of the posteriors deviating by about $-4.6\,\sigma$ and $-3.5\,\sigma$, respectively. This is a result of the assumption that Belle II confirms the current central values measured by the LHCb experiment, in particular, those of the B anomalies in

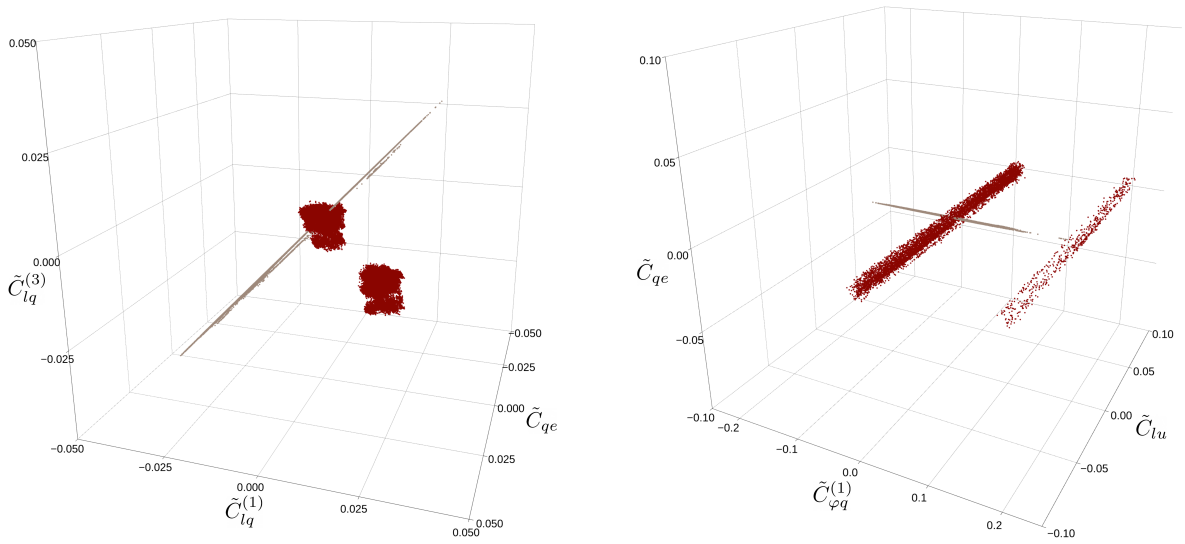

Figure 8.13: Three-dimensional posterior distributions of $\tilde{C}_{lq}^{(1)}$ vs. $\tilde{C}_{lq}^{(3)}$ vs. $\tilde{C}_{qe}$ (left) and $\tilde{C}_{\varphi q}^{(1)}$ vs. $\tilde{C}_{qe}$ vs. $\tilde{C}_{lu}$ (right) obtained in fits of the eleven coefficients in Fig. 8.1 to the combined current and near-future measurements (red), to only the CLIC measurements (light brown), and to the total combination of all measurements (dark brown).[3]

$B^0 \to K^* \mu^+ \mu^-$. While this assumption was also made for the near-future fits of Sec. 8.4.1, the precision of the constraints did not allow observing such deviations in the corresponding fits.

In combination with the expected CLIC data, however, the constraints are strong enough to observe the deviations. If instead assuming the current SM predictions as the nominal values for the Belle II projections of the $B^0 \to K^* \mu^+ \mu^-$ measurements, the constraints of all coefficients contain the SM value within the 90 % HDRs. This is visualized for $\tilde{C}_{lq}^{(1)}$ and $\tilde{C}_{lq}^{(3)}$ in Fig. 8.14, where the two-dimensional 90 % HDRs and the one-dimensional marginal distributions obtained in the total combination of current, near-future, and far-future measurements are shown when assuming that Belle II confirms the B anomalies measured by LHCb (left) and when assuming Belle II measures the corresponding SM values (right).

[3] The posterior distributions of the total combination in the dark brown color are also plotted but cannot really be seen because they are very small and covered by the other distributions.

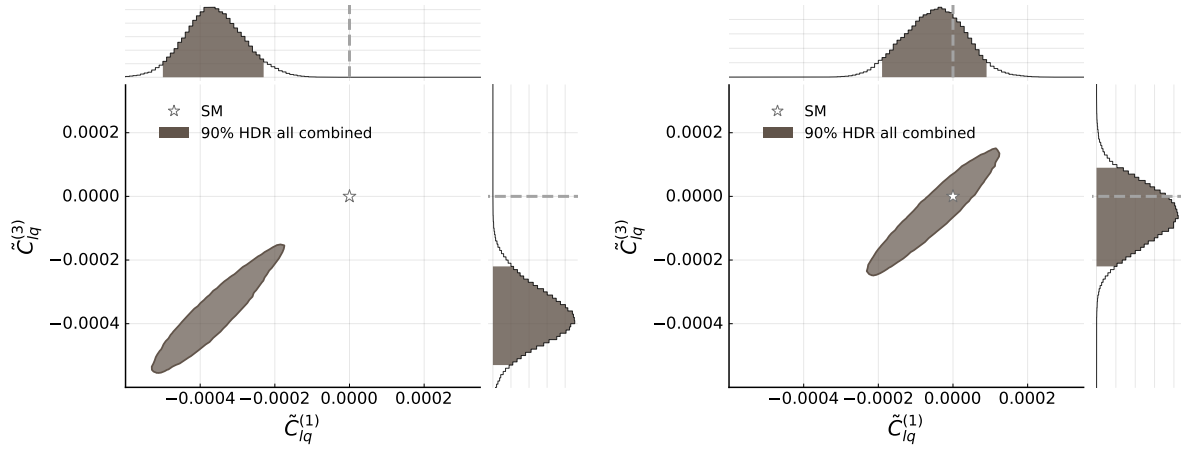

Figure 8.14: One- and two-dimensional 90 % HDRs of the posterior distributions obtained in fits of the eleven coefficients in Fig. 8.1 to the combined current, near-future, and far-future measurements when assuming Belle II confirms the observed B anomalies in the $B^0 \to K^*\mu^+\mu^-$ measurements (left) and when assuming Belle II measures the corresponding SM predictions (right). The SM value is indicated.

8.5 Sampling multimodal distributions with BAT.jl

As observed in Sec. 8.3.3 and Sec. 8.4, the posterior distributions obtained in the combined SMEFT fits of up to eleven coefficients are multimodal. It is a known drawback of classical MCMC approaches, such as the Metropolis–Hastings algorithm or HMC, that these methods are not well suited for sampling multimodal target distributions [256, 257]. This is because the Markov chains constructed with these algorithms only rarely jump between the modes of a distribution, in particular, if the modes are isolated and have very different heights. In the SMEFT fits above, this complicates the exploration of the multimodal posterior distributions with the default MCMC algorithms of BAT.jl. Reliably sampling the corresponding distributions is only possible using the nested sampling algorithms introduced in Sec. 3.3.5. In the following, the difficulties in sampling multimodal distributions with the default Metropolis–Hastings algorithm of BAT.jl are demonstrated using a low-dimensional SMEFT fit as an example.

For this purpose, a fit of the four coefficients $\tilde{C}_{uB}$, $\tilde{C}_{\varphi q}^{(1)}$, $\tilde{C}_{qe}$, and $\tilde{C}_{lq}^{+}$ to the combination of the current top-quark, B, and $Z \to b\bar{b}$ data sets is performed. Uniform prior distributions with maximum ranges of $|\tilde{C}_i| \leq 1$ are chosen for all parameters. It is, thus, the same setup as for the combined fit of Fig. 8.6, but considering only four SMEFT coefficients instead of eight. As introduced in Sec. 3.3.1, for such low-dimensional cases, BAT.jl provides algorithms that evaluate the target distributions at a number of sampling points in the parameter space that are not adapted to the shape of the distribution. These algorithms always yield independent samples of the true target distributions and do not rely on any conditions, such as the convergence of iterative computations. In this example, the Sobol algorithm is used for exploring the posterior distribution, and the corresponding results are depicted in the top plot of Fig. 8.15 in terms of the one- and two-dimensional marginal distributions.

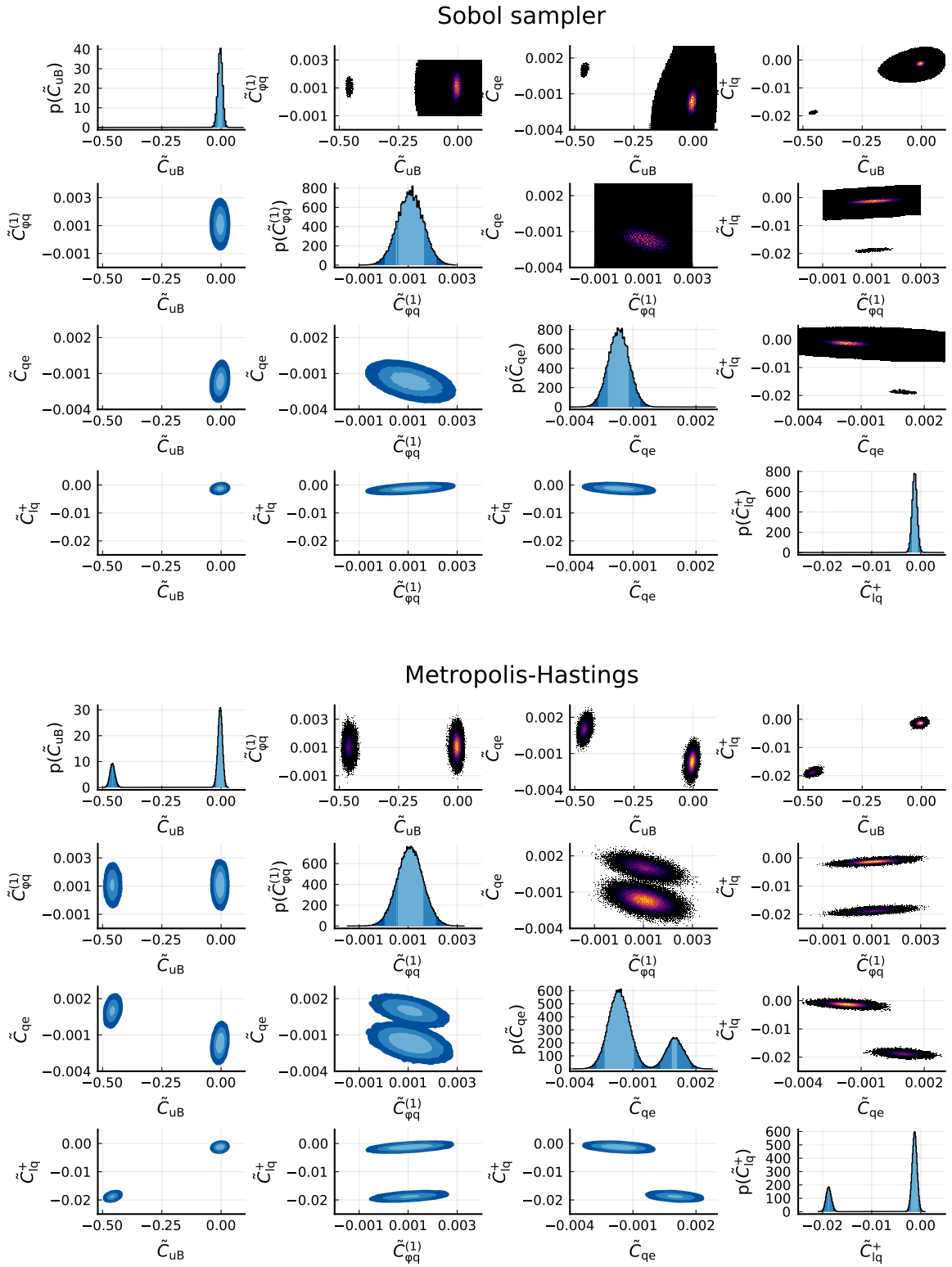

Figure 8.15: Posterior distributions obtained in fits of the four coefficients $\tilde{C}_{uB}$, $\tilde{C}^{(1)}_{\varphi q}$, $\tilde{C}_{qe}$, and $\tilde{C}^{+}_{lq}$ to the combination of current top-quark, B, and $Z \to b\bar{b}$ physics data sets. One- and two-dimensional marginalized distributions are shown for the posterior distributions obtained with the Sobol sampler (top) and with the Metropolis–Hastings algorithm (bottom). The 68.3 %, 95.5 %, and 99.7 % HDRs are highlighted.

The plots in the upper triangle show two-dimensional heatmaps of the posterior probability distribution, and the plots in the lower triangle depict the 68.3 %, 95.5 %, and 99.7 % HDRs. Considering the one-dimensional marginal distributions and the HDRs, it is observed that for each of the SMEFT coefficients, a unimodal distribution is obtained. Only when taking a closer look at the heatmaps in the upper triangle, the presence of multiple modes for $\tilde{C}_{uB}$, $\tilde{C}_{qe}$, and $\tilde{C}_{lq}^{+}$ can be suspected. However, the relative importance of the additional modes is negligibly small, as they are not observed in the one-dimensional marginal distributions and are not part of the considered HDRs.

In Fig. 8.16, the one-dimensional marginal distributions of $\tilde{C}_{uB}$ (left) and $\tilde{C}_{lq}^{+}$ (right), obtained with the Sobol sampler, are examined in more detail.

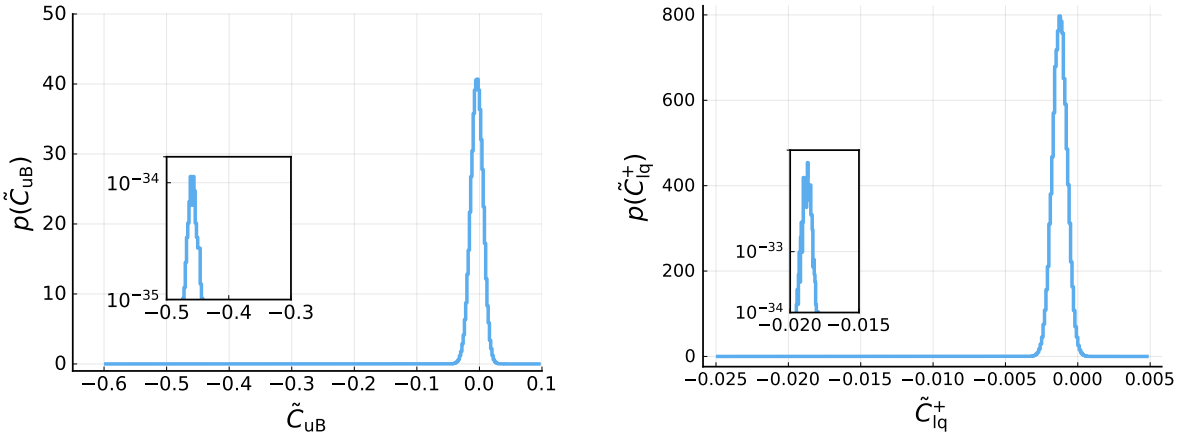

Figure 8.16: One-dimensional marginalized distributions of $\tilde{C}_{uB}$ (left) and $\tilde{C}_{lq}^{+}$ (right) obtained using the Sobol sampler.

In the inset plots, the y-axes are zoomed in, and second modes of the posterior distributions are indeed observed. However, these modes are many orders of magnitude smaller than the main peaks. This reveals that the four-dimensional posterior distribution of the considered example has one global mode but additional strongly suppressed local modes.

When using the Metropolis–Hastings algorithm for sampling this posterior distribution, varying results are obtained when repeating the fit multiple times. On the one hand, unimodal distributions similar to the top plot of Fig. 8.15 are observed. On the other hand, multimodal distributions with varying heights of the individual modes are obtained. In the bottom plot of Fig. 8.15, one example of such a posterior distribution sampled with the Metropolis–Hastings algorithm is shown. Bimodal marginal distributions are observed for the coefficients $\tilde{C}_{uB}$, $\tilde{C}_{qe}$, and $\tilde{C}_{lq}^{+}$, with the second modes taking a significant amount of the posterior probability and being part of the considered HDRs. Two further examples of posterior distributions obtained in repeated fits with the Metropolis–Hastings algorithm are shown in Fig. E.2 of App. E. There, the varying heights and different numbers of modes discovered are visible. These plots reveal that when using the Metropolis–Hastings algorithm for sampling, the relative importance of the local modes is vastly overestimated, if they are discovered at all. This is due to the fact that the Markov chains constructed with the Metropolis–Hastings algorithms typi-

cally only explore the phase space around the mode in which they started and rarely jump between different modes. This leads to an incorrect estimation of the mode heights since in the classical Metropolis–Hastings algorithm, there is no mechanism for weighting the samples from the individual chains according to the probability they covered.

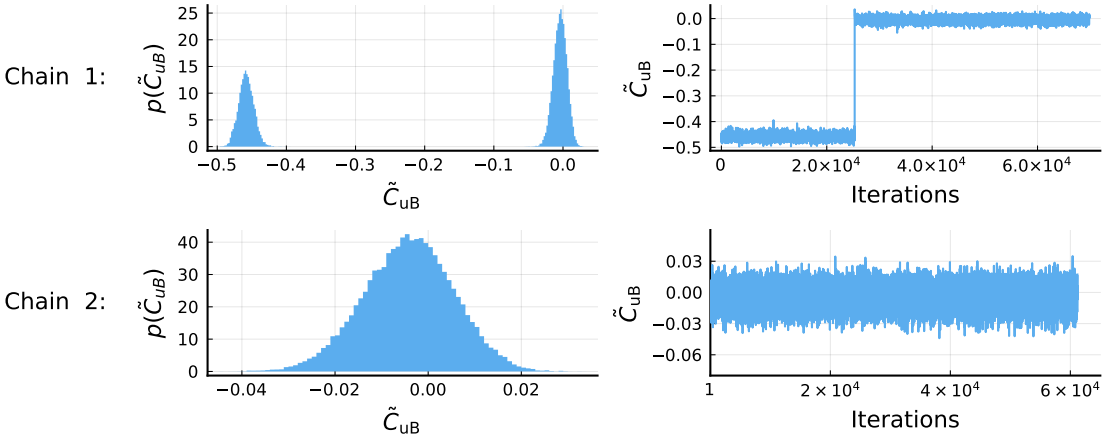

Figure 8.17: Marginal distributions of $\tilde{C}_{uB}$ (left) and trace plots (right) of two individual Markov chains when sampling the four-dimensional examples with the Metropolis–Hastings algorithm.

In Fig. 8.17, the behavior of two Markov chains sampling the considered four-dimensional posterior distribution is visualized. In the left plots, the marginal distributions of $\tilde{C}_{uB}$ are shown for the individual chains. On the right-hand side, the corresponding trace plots are given, i.e., the chronological sequence of states visited by the respective chains. While in the lower plots, i.e., for *Chain 2*, only the global mode around $\tilde{C}_{uB} \approx 0$ is discovered, *Chain 1* initially starts close to the local mode around $\tilde{C}_{uB} \approx -0.45$. The chain explores this region around the local mode for about 2.5×10^4 iterations before eventually jumping to the global mode and exploring this mode for further 4.5×10^4 iterations. However, such jumps from one mode to another are very rare and typically happen from the smaller modes into the larger ones, as the tails of the larger mode attract the chain. The farther the modes are separated, the less likely such transitions become, and the Markov chains exhibit the very slow mixing observed for *Chain 1*.

In BAT.jl, the convergence criteria introduced in Sec. 3.3.2 are used to determine whether the Markov chains have converged to their stationary distribution or not. For the samples of the Metropolis–Hastings algorithm shown in the lower plot of Fig. 8.15, these criteria indicate that the chains have not converged. When considering Fig. 8.17, this is apparent as the within-chain variance and the between-chain variance, used for determining the convergence, obviously do not agree, as would be required for convergence (see Sec. 3.3.2). If, however, by pure chance, all Markov chains start in the same (local or global) mode and stay there during the whole sampling process, the convergence criteria indicate convergence, even though the other modes of the distribution have never been visited by any chain. This is, for example, observed in the top plot of Fig. E.2, where all Markov chains only sampled the global mode, and no other modes are observed in the two-dimensional heatmaps of the upper triangle. This points out

that using only the Metropolis–Hastings algorithm and the corresponding convergence criteria does not always allow deciding whether the posterior distribution has been sampled correctly. In cases where the shape of the target distribution is not known before and multiple modes might be possible, the stability of the results from the Metropolis–Hastings algorithm should be checked by performing multiple runs and using a large number of chains with different starting points. When an unstable behavior is observed, as in the example above, or when a multimodal posterior distribution is already expected, other sampling methods should be used for exploring the distribution and performing cross-checks.

There are approaches for constructing adaptive MCMC algorithms that allow for more efficient exploration of multimodal distributions [256–258]. In BAT.jl, different algorithms and approaches are available for sampling such multimodal target distributions. As mentioned before, for low-dimensional problems, BAT.jl offers the algorithms of Sec. 3.3.1, which always yield the correct distributions. For problems with more than five parameters, the nested sampling algorithms introduced in Sec. 3.3.5 are the best choice for exploring multimodal distributions. Since in nested sampling, the phase space is explored using multiple live points that converge towards the modes and the evidence integral is taken into account, the exploration of local and global modes, as well as a correct estimation of their heights, is ensured. For the SMEFT fits of this chapter, both nested sampling algorithms available in BAT.jl (see Sec. 3.3.5) have been applied. They performed well in sampling the multimodal distributions with up to eleven parameters and generated agreeing samples of the considered posterior distributions. Another approach for sampling multimodal distributions in BAT.jl is provided in the form of the partitioned sampling, introduced in Sec. 3.3.4. If the partitions of the parameter space created by this algorithm match the modes of the target distribution, the classical MCMC algorithms can be used for sampling since the Markov chains then only need to explore one mode in each partition. Due to the subsequent AHMI integration when combining the samples from all subspaces, the correct weighting of the probability covered by the individual modes is ensured. In Fig. E.1 of App. E, a posterior distribution of the four-dimensional example, obtained with the partitioned sampling approach and the Metropolis–Hastings algorithm, is shown. There, it is observed (in the heatmaps of the upper triangle) that the local modes are discovered. However, they correctly receive small weights and are neither visible in the one-dimensional distributions nor part of the HDRs. While in the low-dimensional example the partitioned sampling approach performed comparably well as the nested sampling algorithm, for the SMEFT fits in Sec. 8.4 with eleven parameters and multiple modes, the nested sampling algorithms provided more reliable results and required less manual tuning of hyperparameters.

This example demonstrates the complications of exploring (potentially) multimodal posterior distributions with the Metropolis–Hastings algorithm. It points out that even in cases where the target distribution can be considered unimodal for all practical purposes, the presence of additional (but strongly suppressed) modes can lead to wrong results of the sampling. This also stresses the importance of choosing appropriate sampling algorithms for the problem at hand and highlights that performing cross-checks using different techniques is an essential aspect, in particular when no prior knowledge about the shape of the posterior distribution is

available. In BAT.jl, the collection of different sampling algorithms facilitates performing such cross-checks and makes it simple to choose the best-suited algorithm for the current problem.

8.6 Summary

In the studies above, the approach of Chap. 7 was extended, and further observables and SMEFT coefficients were included in a combined fit. It was demonstrated again that a combination of observables from different energy scales can lead to a significant improvement of the constraints on the SMEFT coefficients. By combining current measurements of top-quark production and decay processes, of $b \to s$ transitions, and of $Z \to b\bar{b}$ decays, it was illustrated how the different sensitivities of the corresponding observables can yield orthogonal constraints on the Wilson coefficients. In particular, it was pointed out that such complementary sensitivities on certain coefficients also affect the multidimensional parameter space. This allows improving the constraints on additional coefficients to which the observables themselves are not directly sensitive. By including projections of measurements expected in the near and far future, the physics potential of the considered future experiments was pointed out. It was demonstrated that the improved precision of the anticipated measurements can lead to tighter constraints and allows constraining more SMEFT coefficients. Specifically, it was shown that expected observations of $b \to s\nu\bar{\nu}$ transitions at Belle II allow resolving flat directions in the constraints and enable probing the coefficients $\tilde{C}_{lq}^{(1)}$ and $\tilde{C}_{lq}^{(3)}$ independently. The studies assuming a CLIC-like lepton collider demonstrated the significant improvements of the constraints on the considered SMEFT coefficients to be expected from such a future collider. Again, the benefits of combining observables with complementary sensitivities for resolving ambiguous solutions were highlighted. Performing these SMEFT fits with up to eleven free parameters also demonstrated the capabilities of the new implementations of EFTfitter.jl and BAT.jl. The posterior distributions obtained in the combined SMEFT fits are multimodal and have complicated shapes in the multidimensional phase space. It was pointed out in Sec. 8.5 that classical MCMC methods are not suited for such problems and that only the nested sampling algorithms available in BAT.jl are capable of reliably sampling these posterior distributions. As these algorithms were not available in the predecessor versions of the tools, these studies also demonstrated the excellent suitability of the new BAT.jl and EFTfitter.jl packages for applications in SMEFT interpretations, where such multimodal posterior distributions are common, due to the often quadratic dependences of the observables on the Wilson coefficients.

9 Conclusion

Bayesian inference provides a systematic mathematical approach to gaining knowledge by comparing model expectations with observed data. While interpretations of the results are straightforward, applying Bayesian inference for data analysis typically requires efficient numerical algorithms for sampling and integration. In this thesis, the newly developed *Bayesian Analysis Toolkit in Julia* (BAT.jl) was introduced. Its design concepts were motivated, and the included collection of state-of-the-art algorithms for posterior exploration and integration was presented. As direct searches for *physics beyond the Standard Model* (BSM) have not yet led to any observations of new particles, *effective field theories* (EFTs) provide an indirect approach to probing BSM physics at energy scales much higher than directly accessible with current collider experiments. Interpretations of measurements in the context of the *Standard Model effective field theory* (SMEFT) gained much interest in recent years, particularly in the top-quark sector. In this thesis, methods for Bayesian data analysis were applied in BSM searches using the SMEFT formalism. For this purpose, the newly revised EFTfitter.jl package was introduced in Chap. 5 as a tool facilitating EFT interpretations of measurements. Its concepts for combining measurements of different observables and constraining underlying parameters with Bayesian inference were discussed. In particular, methods for assigning weights to the individual measurements describing their impact on the total combination were studied and compared. It was pointed out that weights based on the sizes of highest density regions are a valid approach for arbitrarily shaped posterior distributions but can currently not be employed reliably for problems with more than three parameters.

The tools BAT.jl and EFTfitter.jl were applied in this thesis to constrain Wilson coefficients of dimension-six SMEFT operators. Particular focus was placed on investigating approaches for enhancing SMEFT fits in the top-quark sector. It was discussed that combinations of measurements are a common method for constraining SMEFT coefficients, but correlations between the uncertainties of the measurements are often not known precisely and, as a consequence, frequently neglected in such analyses. Therefore, studies on the impact of correlations on the results of SMEFT fits were performed in Chap. 6. Observables of top-quark production and decay processes were considered, and different correlation scenarios were constructed. A *best guess correlations* scenario was developed using simplifying assumptions and compared to the *known correlations* scenario for quantitatively investigating the effects of correlations on the resulting constraints. The studies revealed that neglecting correlations can lead to significantly different results and demonstrated that correlations can be the crucial components of SMEFT fits deciding whether new phenomena are discovered or not. By considering projections for future measurements with reduced uncertainties, as expected from the HL-LHC, it was pointed out that the impact of correlations will become even more critical in the future. These studies thus illustrated that special attention needs to be paid to the treatment of correlations, particularly when aiming towards global fits of SMEFT coefficients. For such fits, the analyses of Chap. 6 can serve as a guideline for investigating the impact of correlations, and similar studies could be employed to validate the results of SMEFT fits with potentially unknown correlations between the uncertainties of measurements.

Further enhancements of SMEFT fits were pursued by conducting first studies on the combination of measurements from top-quark and B physics in Chap. 7. The steps necessary for a combined fit of observables from different energy scales were laid out, and Wilson coefficients of operators affecting the top-quark couplings were constrained. By comparing the results of fits to the individual observables with the results of the combined fit, significant improvements of the constraints by more than one order of magnitude were highlighted as a result of complementary sensitivities. The observed synergies between top-quark and flavor physics were examined further in Chap. 8 by extending the analyses to include more Wilson coefficients, observables, and measurements. SMEFT fits using data from top-quark production and decay, $Z \to b\bar{b}$ decay, and $b \to s$ transitions were performed, and up to eleven SMEFT coefficients were constrained. Orthogonal constraints in the multidimensional parameter spaces originating from the different types of observables were discovered, and the benefits of combining them to tighten the constraints and resolve ambiguous solutions were illustrated. By performing studies of future scenarios, including projections for measurements expected from the HL-LHC, Belle II, and the CLIC lepton collider, the potential benefits of these experiments for constraining BSM physics were pointed out. In particular, it was demonstrated that the considered future measurements could allow deriving tight constraints on currently weakly constrained SMEFT coefficients. These studies also highlighted the advantages of the comprehensive collection of algorithms provided by the new implementation of BAT.jl. It was demonstrated that in some of the high-dimensional SMEFT fits, multimodal posterior distributions were obtained, and only nested sampling algorithms were suited to sample these distributions reliably.

In conclusion, the studies presented in this thesis served, on the one hand, as first use cases demonstrating the capabilities of the newly implemented BAT.jl and EFTfitter.jl packages for BSM searches. On the other hand, they pointed out important aspects for enhancing SMEFT fits, which should be taken into account when pursuing the goal of global fits to current and future data.

Appendices

Appendix A

BAT.jl default output

```
BAT.jl - SampledDensity
──────────────────────────────


Sampling:
──────────────────────────
total number of samples:      1560112
effective number of samples:
(a = 322005.24040657695, b = 243304.36878302894,
 c = 315216.9656156388)


Parameter estimates:
──────────────────────────
number of free parameters: 3

Table with 5 columns and 3 rows:
     parameter  mean          std       global_mode   marginal_mode
   ┌──────────────────────────────────────────────────────────────────
 1 │ a          -0.000905671  0.998413  -0.000331011  -0.075
 2 │ b          1.99818       4.90005   5.00202       4.9
 3 │ c          4.99551       1.99696   5.02656       4.85


Covariance matrix:
──────────────────────────
3×3 Named Array{Float64,2}
cov \  │           a            b            c
──────┼──────────────────────────────────────────
a      │    0.996829   0.00662992  0.000480882
b      │  0.00662992      24.0105   -0.0146034
c      │ 0.000480882   -0.0146034      3.98785
```

Figure A.1: Example output obtained from a BAT.jl `SampledDensity` object. The output shown here corresponds to the posterior distribution visualized in Fig. 3.1.

Appendix B

Additional plots on the weight studies

In this appendix, additional plots related to the studies on the influence of individual measurements on their combination in Sec. 5.4 are given.

B.1 Correlation matrix

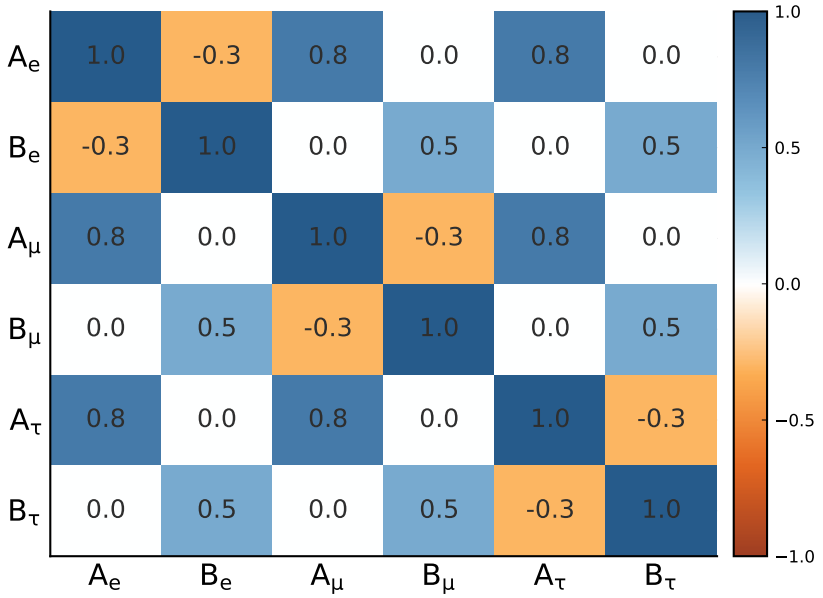

Figure B.1: Correlation matrix assumed for the W boson decay measurements in the studies of Sec. 5.4

B.2 Weights in the non-lepton-universal example

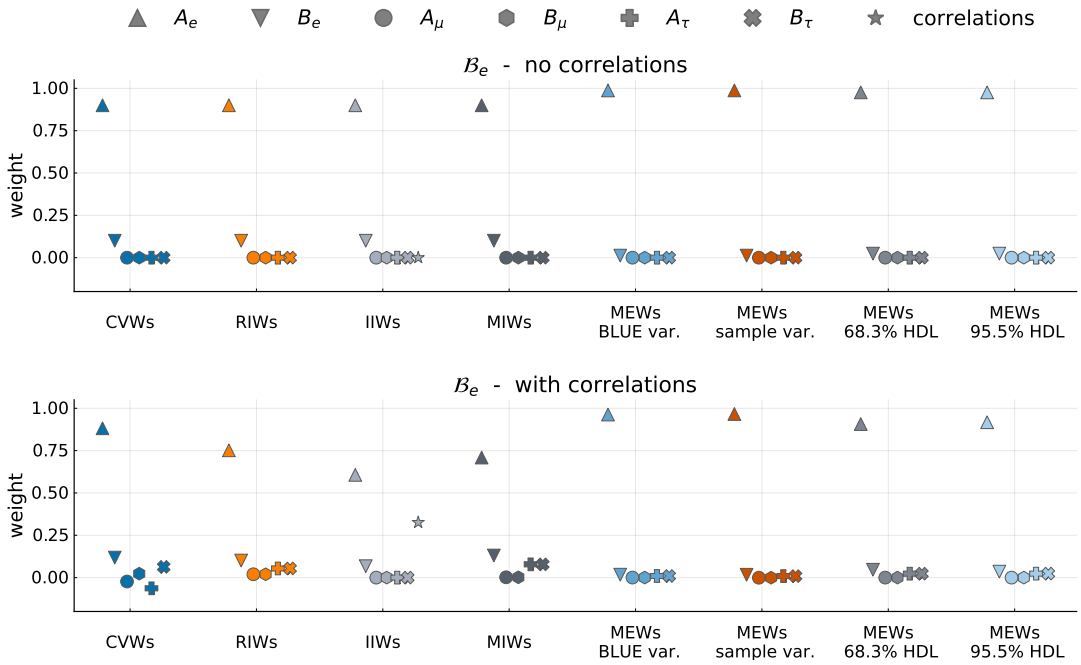

Figure B.2: Comparison of the different types of weights for the electron channel when assuming no correlations between the measurements (top) and when assuming the correlations of Fig. B.1 (bottom).

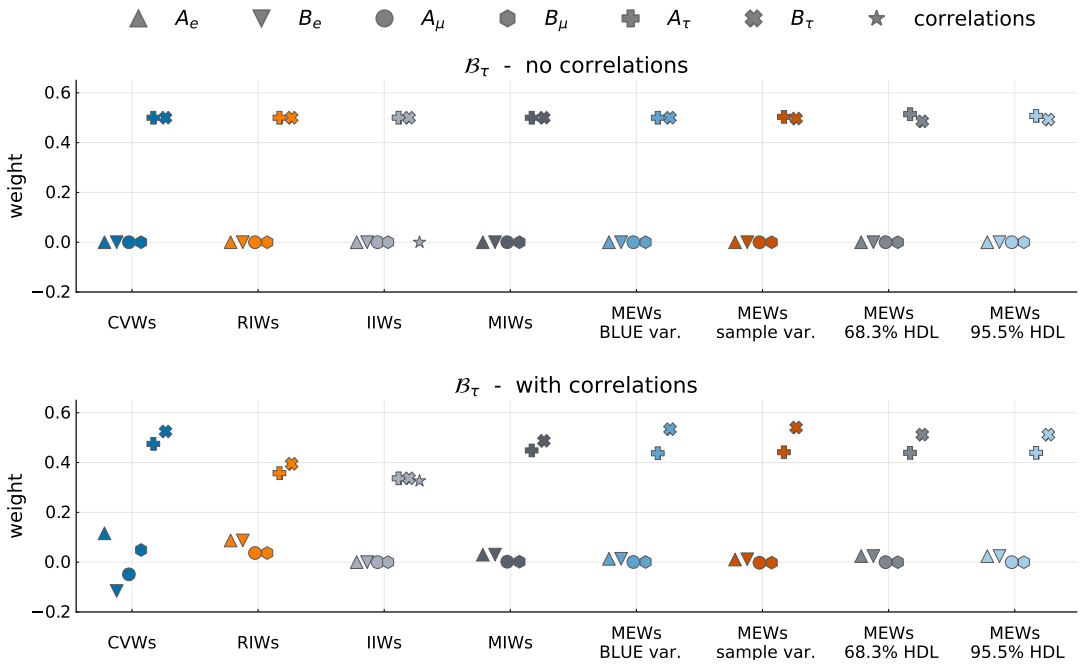

Figure B.3: Comparison of the different types of weights for the tau channel when assuming no correlations between the measurements (top) and when assuming the correlations of Fig. B.1 (bottom).

B.3 Posterior distribution of the example in Sec. 5.4.3 when assuming correlations

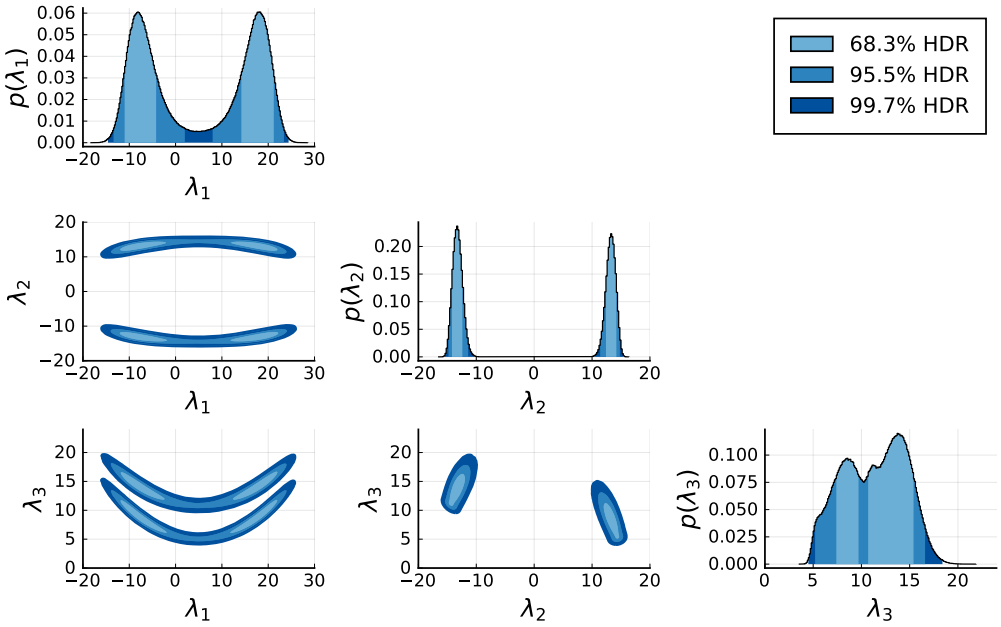

Figure B.4: Posterior distribution of the total combination when assuming the dependence on underlying parameters of Eq. (5.16) and the correlations of Fig. B.1. The one- and two-dimensional marginal distributions are shown, and the HDRs containing 68.3 %, 95.5 %, and 99.7 % of the posterior probability are highlighted.

Appendix C

Additional material on the correlation studies

In this appendix, additional plots and tables related to the studies on the impact of correlations on the results of SMEFT fits in Chap. 6 are given.

C.1 Correlation matrices

The full correlation matrices of the statistical, systematic, and theory uncertainties in the *best guess correlations* scenario are shown in Fig. C.1, Fig. C.2, and Fig. C.3.

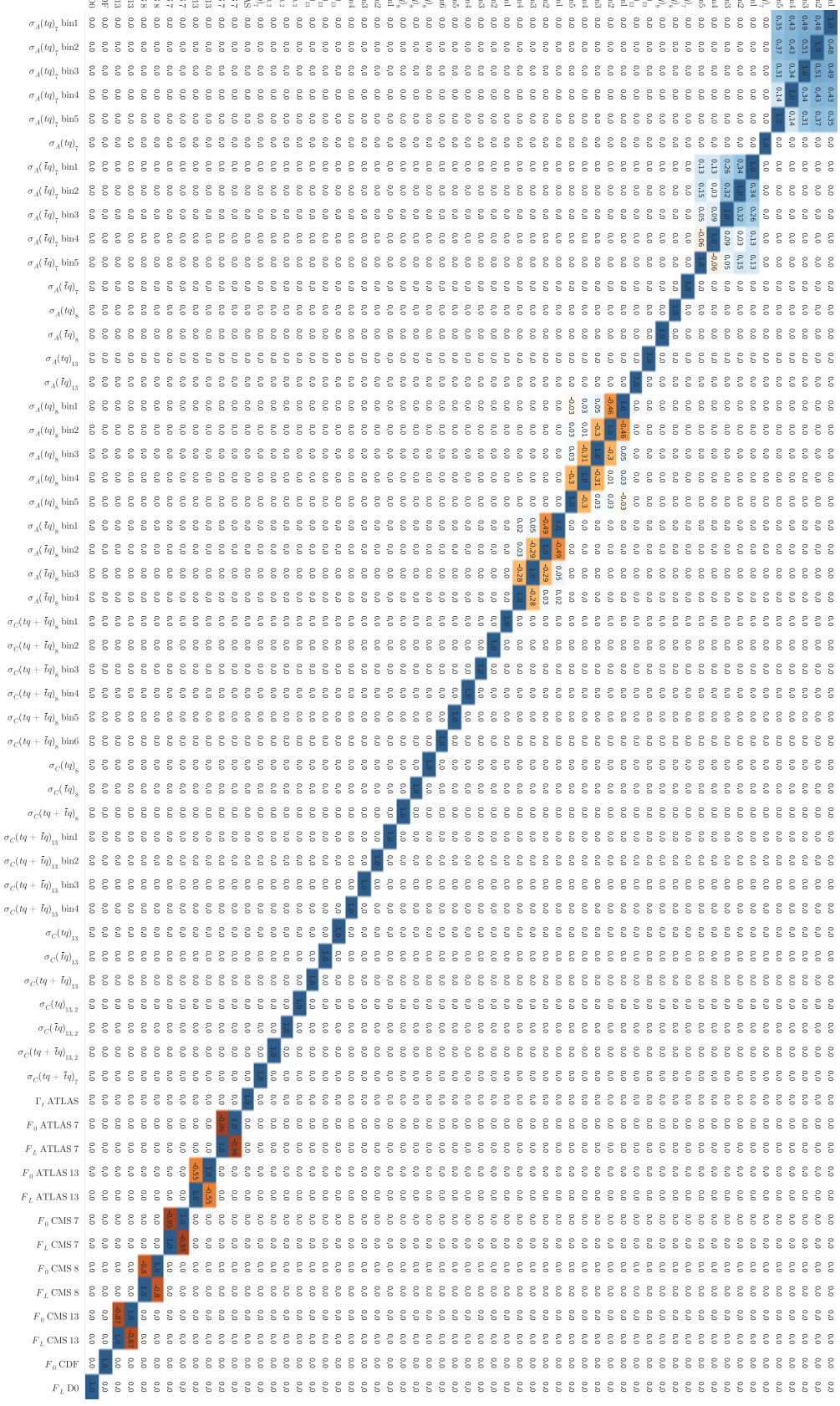

Figure C.1: Full correlation matrix of the systematic uncertainties in both the *known correlations* and the *best guess correlations* scenarios.

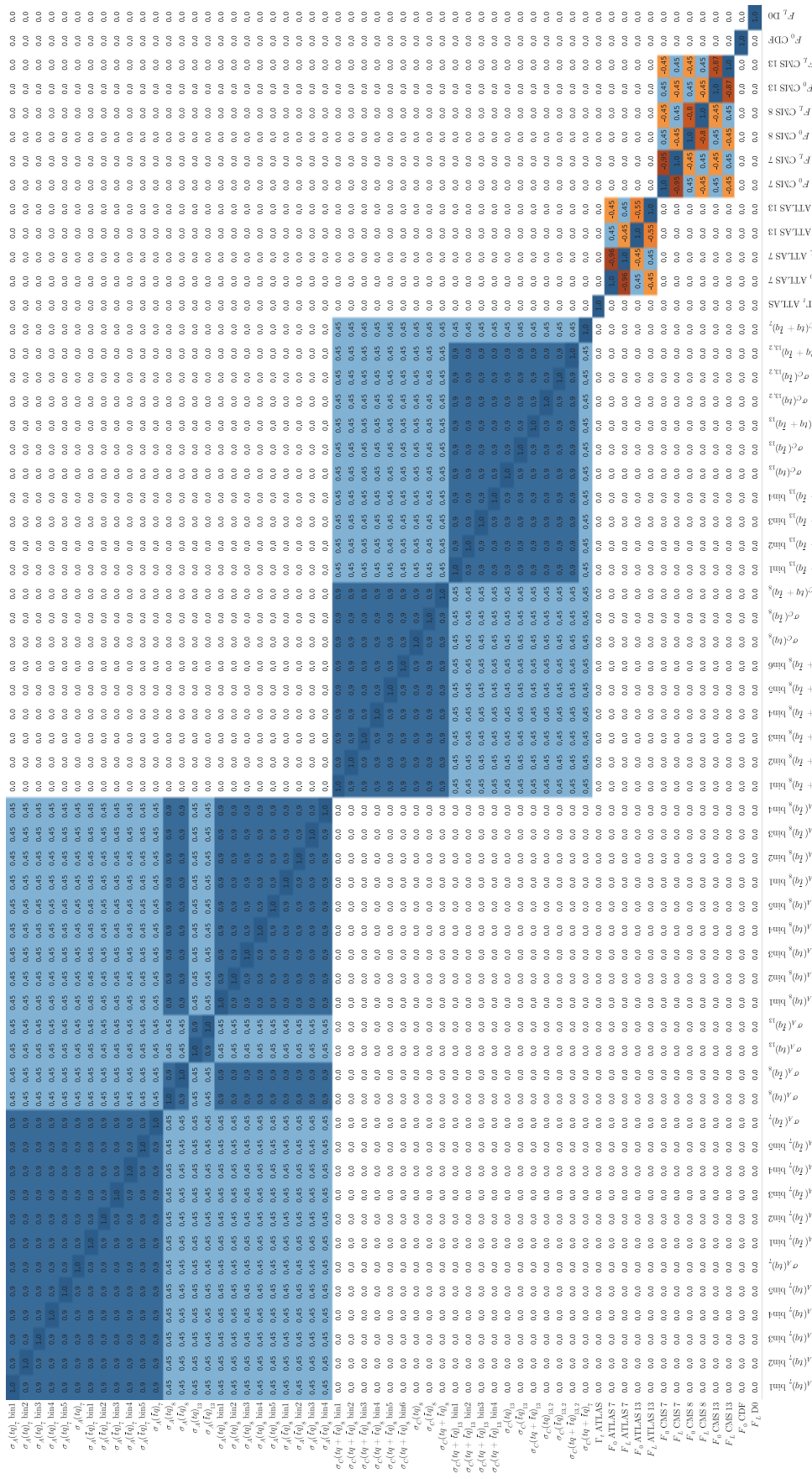

Figure C.2: Full correlation matrix of the systematic uncertainties in the *best guess correlations* scenario.

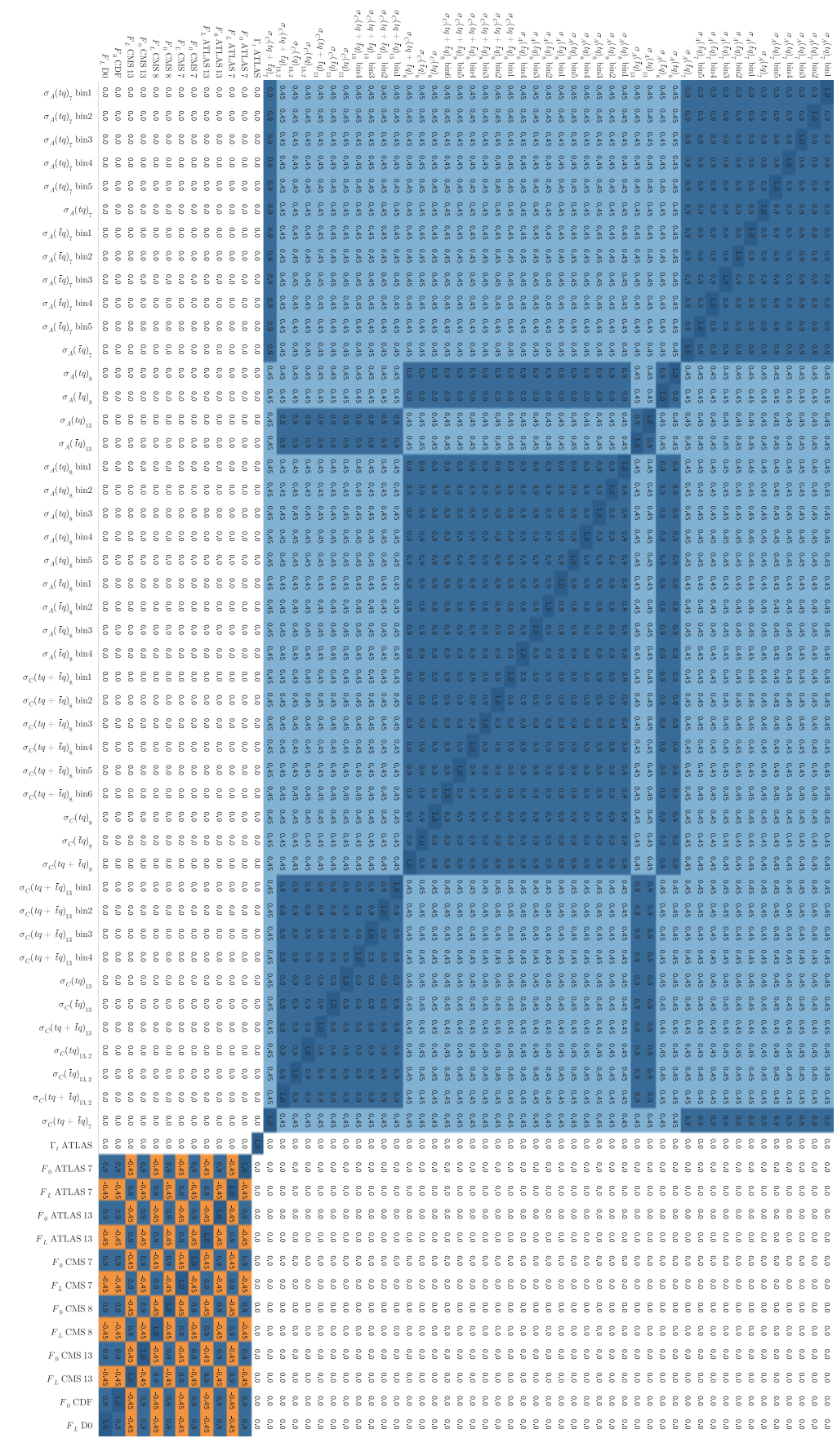

Figure C.3: Full correlation matrix of the theory uncertainties in the *best guess correlations* scenario.

C.2 Posterior distribution of the *best guess correlations* scenario when using the quadratic parametrization

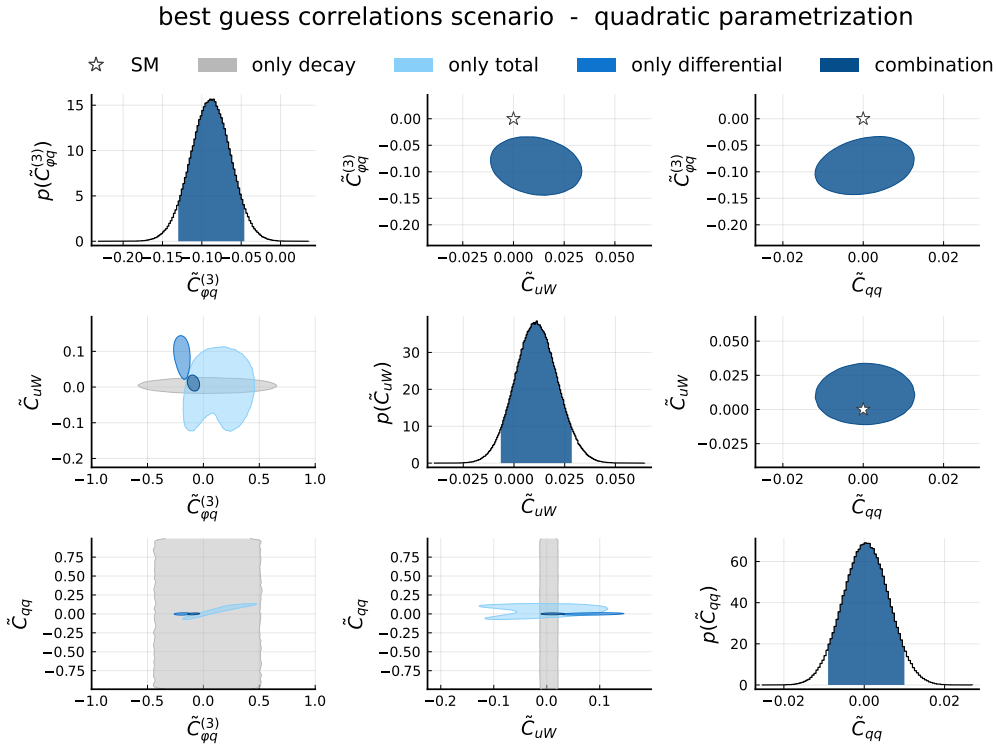

Figure C.4: Posterior distribution of the *best guess correlations* scenario when using the quadratic parametrization of cross sections. Shown are the one- and two-dimensional marginalized posterior regions containing 90 % posterior probability. On the diagonal and in the upper triangle, the results of the total combination of all measurements are shown. In the lower triangle, the 90 % HDAs are shown for fits to the individual types of observables.

C.3 Varying the correlation coefficients

Additional plots on the effects of varying the values of the correlation coefficients in the *best guess correlations* scenario.

C.3.1 Current uncertainties with linear parametrization

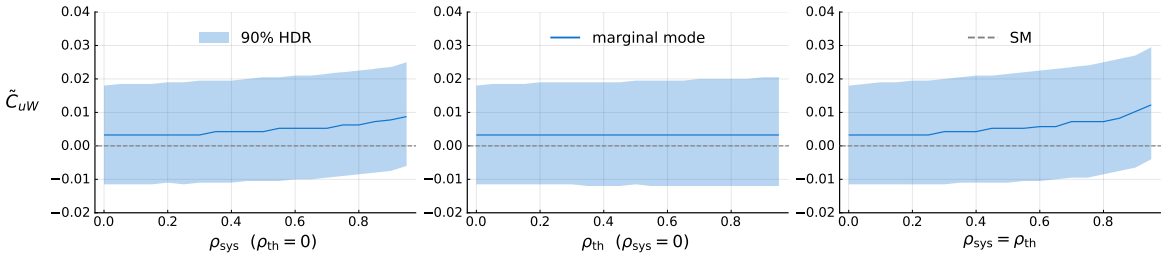

Figure C.5: Marginal mode and smallest 90 % interval for $\tilde{C}_{uW}$ when varying only ρ_{sys} (left), when varying only ρ_{th} (middle), and when varying both simultaneously (right) from 0 to 0.95 in steps of 0.05.

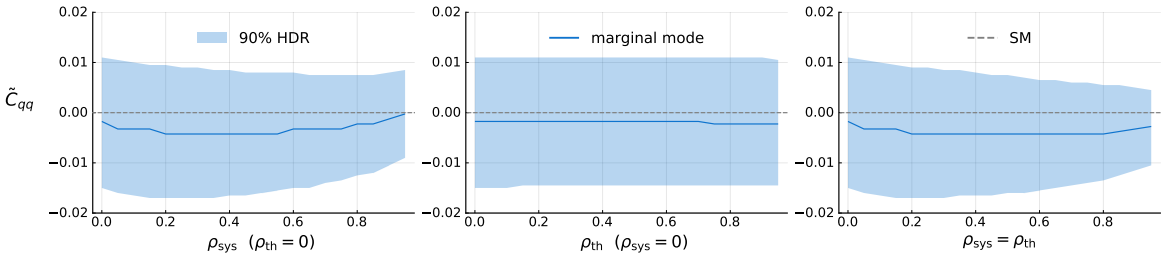

Figure C.6: Marginal mode and smallest 90 % interval for $\tilde{C}_{qq}$ when varying only ρ_{sys} (left), when varying only ρ_{th} (middle), and when varying both simultaneously (right) from 0 to 0.95 in steps of 0.05.

C.3.2 Current uncertainties with quadratic parametrization

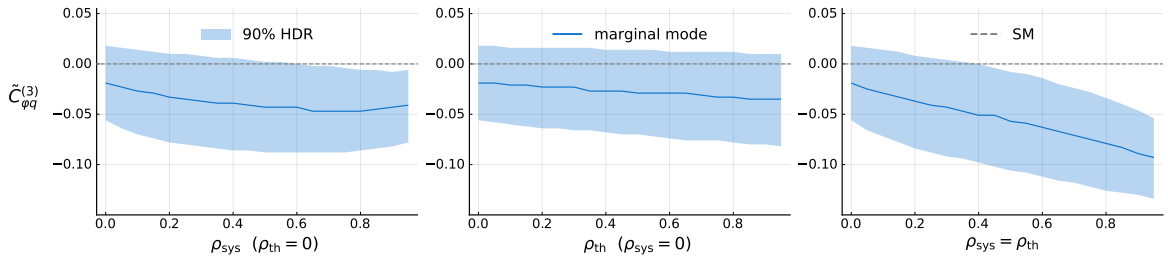

Figure C.7: Marginal mode and smallest 90 % interval for $\tilde{C}^{(3)}_{\varphi q}$ when varying only ρ_{sys} (left), when varying only ρ_{th} (middle), and when varying both simultaneously (right) from 0 to 0.95 in steps of 0.05 using the quadratic parametrization of cross sections.

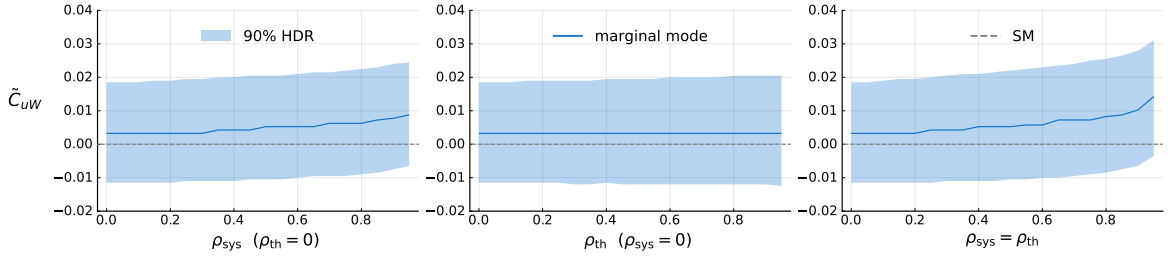

Figure C.8: Marginal mode and smallest 90 % interval for $\tilde{C}_{uW}$ when varying only ρ_{sys} (left), when varying only ρ_{th} (middle), and when varying both simultaneously (right) from 0 to 0.95 in steps of 0.05 using the quadratic parametrization of cross sections.

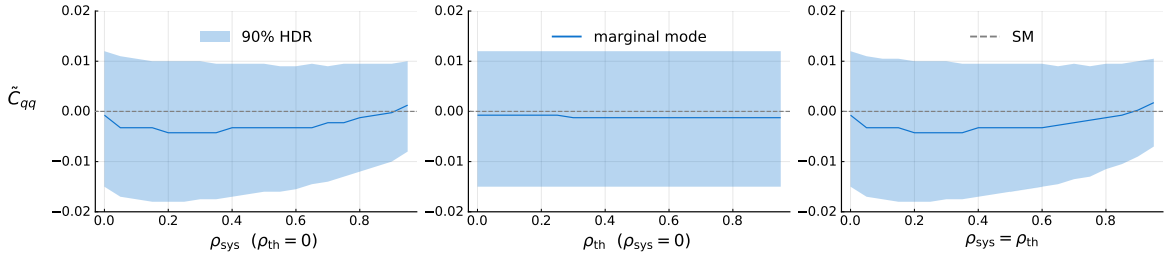

Figure C.9: Marginal mode and smallest 90 % interval for $\tilde{C}_{qq}$ when varying only ρ_{sys} (left), when varying only ρ_{th} (middle), and when varying both simultaneously (right) from 0 to 0.95 in steps of 0.05 using the quadratic parametrization of cross sections.

C.3.3 Current uncertainties with linear parametrization

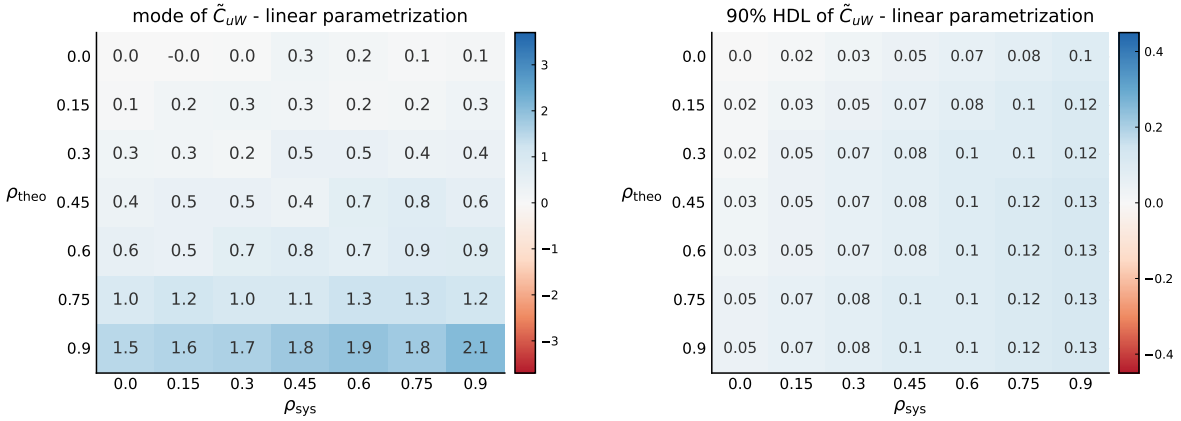

Figure C.10: Relative deviation of the marginal mode (left) and relative increase of the 90 % HDL compared to the *known correlations* scenario for $\tilde{C}_{uW}$ when varying the values of ρ_{sys} and ρ_{th} independently.

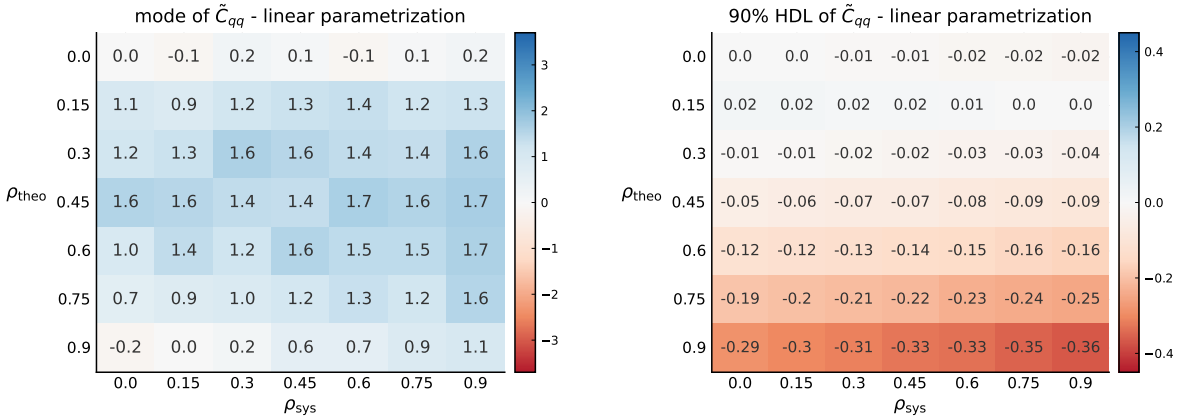

Figure C.11: Relative deviation of the marginal mode (left) and relative increase of the 90 % HDL compared to the *known correlations* scenario for $\tilde{C}_{qq}$ when varying the values of ρ_{sys} and ρ_{th} independently.

C.3.4 Reduced statistical uncertainties with linear parametrization

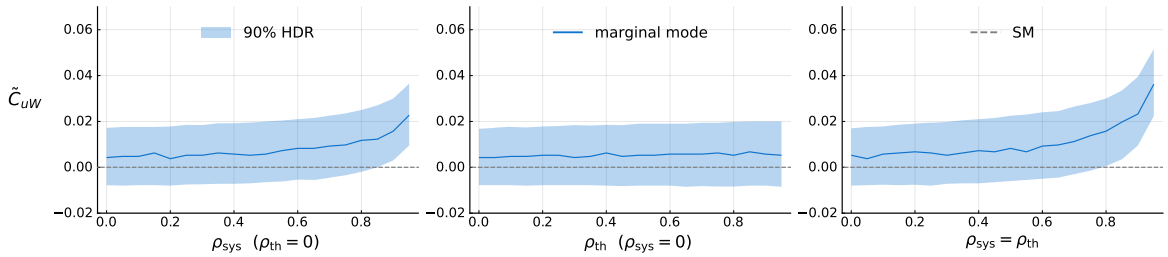

Figure C.12: Marginal mode and smallest 90 % interval for $\tilde{C}_{uW}$ when varying only ρ_{sys} (left), when varying only ρ_{th} (middle), and when varying both simultaneously (right) assuming the reduced statistical uncertainties.

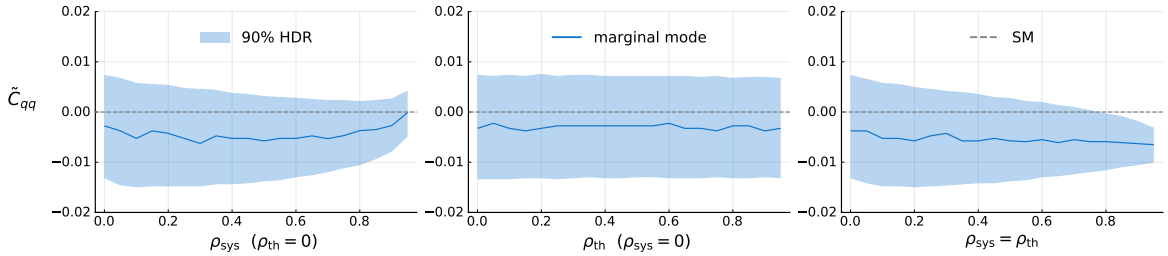

Figure C.13: Marginal mode and smallest 90 % interval for $\tilde{C}_{qq}$ when varying only ρ_{sys} (left), when varying only ρ_{th} (middle), and when varying both simultaneously (right) assuming the reduced statistical uncertainties.

C.4 Reduced statistical, systematic, and theory uncertainties

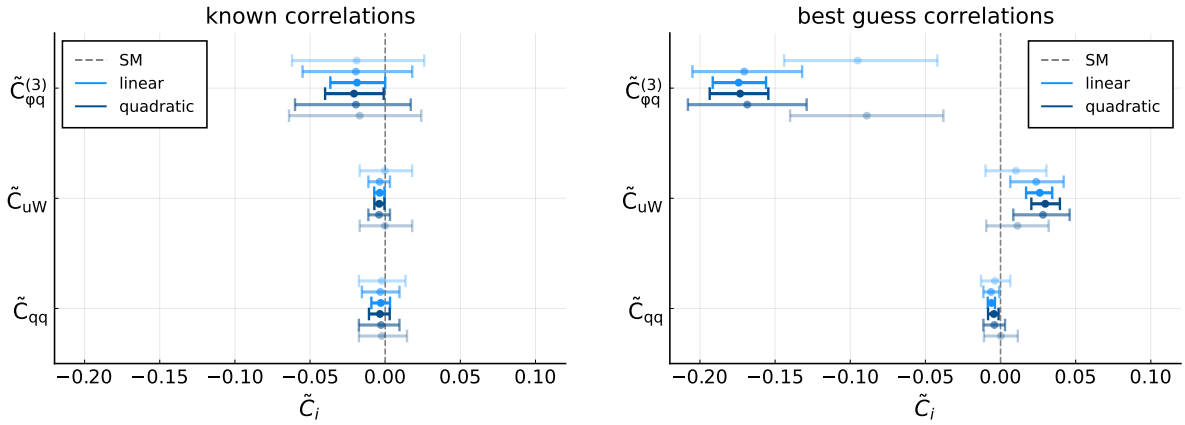

Figure C.14: Marginal modes and smallest 90 %intervals of the one-dimensional marginalized posterior distributions using the linear (light blue) and quadratic (dark blue) parametrization of cross sections in the *known correlations* scenario (left) and in the *best guess correlations* scenario (right) for the statistical uncertainties scaled down to 3000 fb^{-1} and the systematic and theory uncertainties reduced by a factor of two. The SM is indicated as the dashed gray line. The results obtained with the current uncertainties and the reduced statistical uncertainties are indicated using the slightly transparent markers.

Table C.1: Bounds of the marginalized 90 % HDIs, their total length, the standard deviation, and the value of the mode, as well as the deviation of the mode from the SM in units of the standard deviation when assuming the statistical uncertainties reduced to 3000 fb^{-1}, and systematic and theory uncertainties reduced by a factor of two.

	known correlations					*best guess correlations*				
$\tilde{C}_i$	90% HDI	90% HDL	σ	mode	mode/σ	90% HDI	90% HDL	σ	mode	mode/σ
$\tilde{C}^{(3)}_{\varphi q}$	[-0.034, -0.003]	0.031	0.009	−0.019	−1.99	[-0.189, -0.159]	0.030	0.009	−0.174	−19.4
$\tilde{C}_{uW}$	[-0.007, -0.001]	0.006	0.002	0.004	−2.15	[0.018, 0.033]	0.015	0.004	0.026	5.9
$\tilde{C}qq$	[-0.009, 0.002]	0.011	0.003	−0.003	−0.91	[-0.008, -0.004]	0.004	0.001	−0.006	−4.9

Appendix D

Additional material on combinations of top-quark and B physics

D.1 Numerical values of the matching conditions

The numerical values of the tree-level matching conditions of SMEFT and WET at $\mu_W = m_W$ are

$$\Delta C_9^{\text{tree}} = 402.1 \left[\tilde{C}_{lq}^1 + \tilde{C}_{lq}^3 + \tilde{C}_{qe}\right] - 44.53 \left(\tilde{C}_{\varphi q}^1 + \tilde{C}_{\varphi q}^3\right) ,$$

$$\Delta C_{10}^{\text{tree}} = 402.1 \left[-\tilde{C}_{lq}^1 - \tilde{C}_{lq}^3 + \tilde{C}_{qe} + \tilde{C}_{\varphi q}^1 + \tilde{C}_{\varphi q}^3\right] ,$$

$$\Delta C_L^{\text{tree}} = 402.1 \left[\tilde{C}_{lq}^1 - \tilde{C}_{lq}^3 + \tilde{C}_{\varphi q}^1 + \tilde{C}_{\varphi q}^3\right] .$$

The one-loop contributions at $\mu_W = m_W$ are

$$\Delta C_7^{\text{loop}} = -2.310\tilde{C}_{uB} + 0.09251\tilde{C}_{uW} - 0.0946\tilde{C}_{\varphi q}^{(1)} + 0.7951\tilde{C}_{\varphi q}^{(3)} ,$$

$$\Delta C_8^{\text{loop}} = -0.6687\tilde{C}_{uG} + 0.2709\tilde{C}_{uW} + 0.2839\tilde{C}_{\varphi q}^{(1)} + 0.7568\tilde{C}_{\varphi q}^{(3)} ,$$

$$\begin{aligned}\Delta C_9^{\text{loop}} =& 2.170\tilde{C}_{uW} + 2.512\tilde{C}_{uB} + 2.972\tilde{C}_{\varphi q}^{(3)} + 0.2217 \left(\tilde{C}_{\varphi u} - \tilde{C}_{\varphi q}^{(1)}\right) \\ & - 2.002 \left(\tilde{C}_{eu} + \tilde{C}_{lu} - \tilde{C}_{lq}^{(1)} - \tilde{C}_{qe}\right) - 3.174\tilde{C}_{lq}^{(3)} ,\end{aligned}$$

$$\begin{aligned}\Delta C_{10}^{\text{loop}} =& -7.536\tilde{C}_{uW} - 15.70\tilde{C}_{\varphi q}^{(3)} + 4.498 \left(\tilde{C}_{\varphi u} - \tilde{C}_{\varphi q}^{(1)}\right) - 2.002 \left(\tilde{C}_{eu} - \tilde{C}_{lu} + \tilde{C}_{lq}^{(1)}\right. \\ & \left. - \tilde{C}_{qe}\right) + 3.174\tilde{C}_{lq}^{(3)} ,\end{aligned}$$

$$\Delta C_L^{\text{loop}} = -2.876\tilde{C}_{uW} - 5.084\tilde{C}_{\varphi q}^{(3)} - 0.4451 \left(\tilde{C}_{\varphi u} - \tilde{C}_{\varphi q}^{(1)} + \tilde{C}_{lu} - \tilde{C}_{lq}^{(1)}\right) + 0.7057\tilde{C}_{lq}^{(3)} ,$$

$$\Delta C_{1,tt}^{\text{mix, loop}} = 4.120\tilde{C}_{uW} + 10.08\tilde{C}_{\varphi q}^{(3)} .$$

D.2 Fiducial acceptance of the dilepton channel

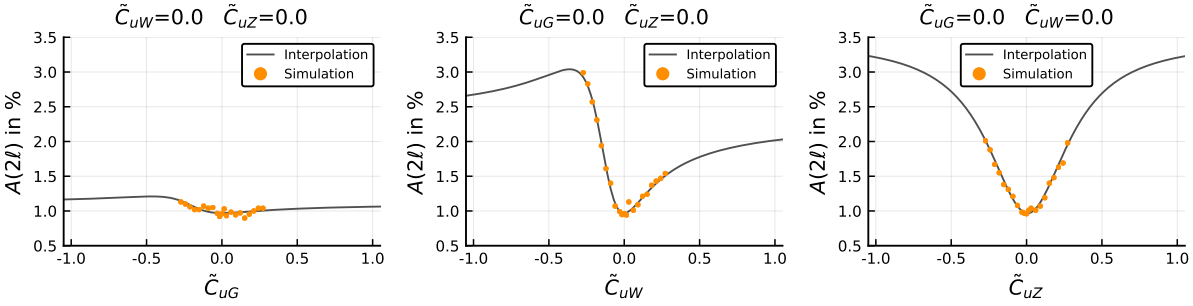

Figure D.1: Sampling points and result of the interpolation for the fiducial acceptance of the dilepton channel $A(2\ell)$. Shown are slices of the phase space where only one of the Wilson coefficient is varied at a time while the others are fixed to zero.

D.3 Posterior distributions of the fit to top-quark measurements

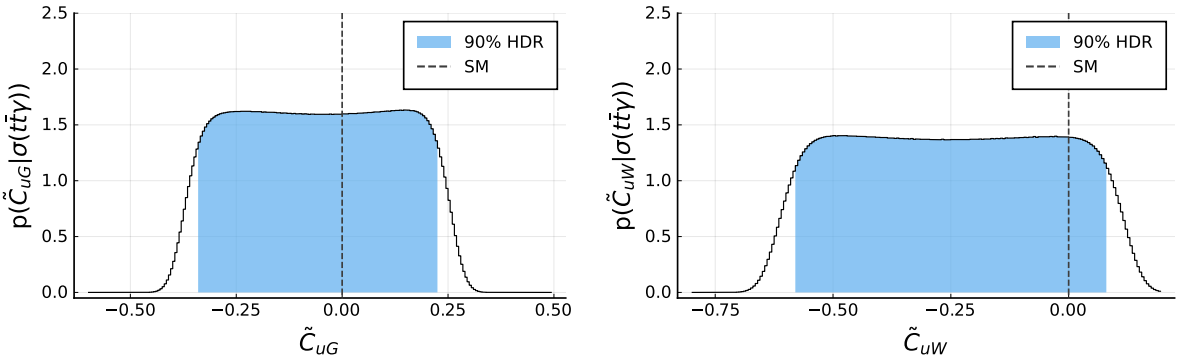

Figure D.2: Marginalized posterior distributions of $\tilde{C}_{uG}$ (left) and $\tilde{C}_{uW}$ (right) when using only the $t\bar{t}\gamma$ cross section measurement. The colored areas highlight the smallest intervals containing 90 % of the posterior probability, and the dashed line indicates the SM value.

D.4 Fit of the WET coefficients to the B measurements

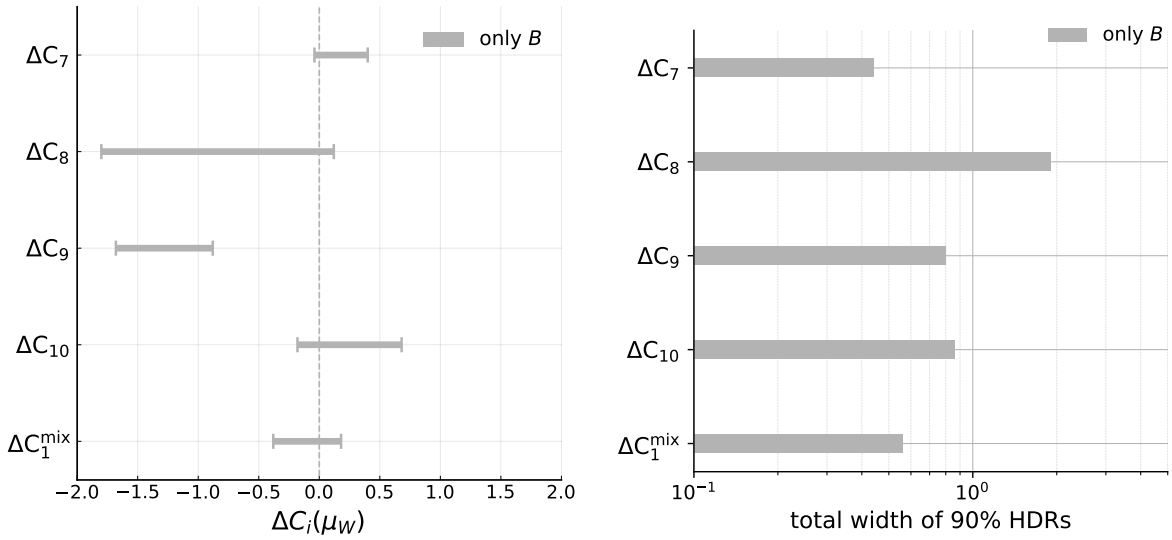

Figure D.3: Marginalized smallest intervals containing 90 % of the posterior probability (left) and their total widths (right) obtained in a fit of the WET coefficients to only the B physics measurements of Tab. 8.3.

Appendix E

Additional plots on the sampling of multimodal distributions

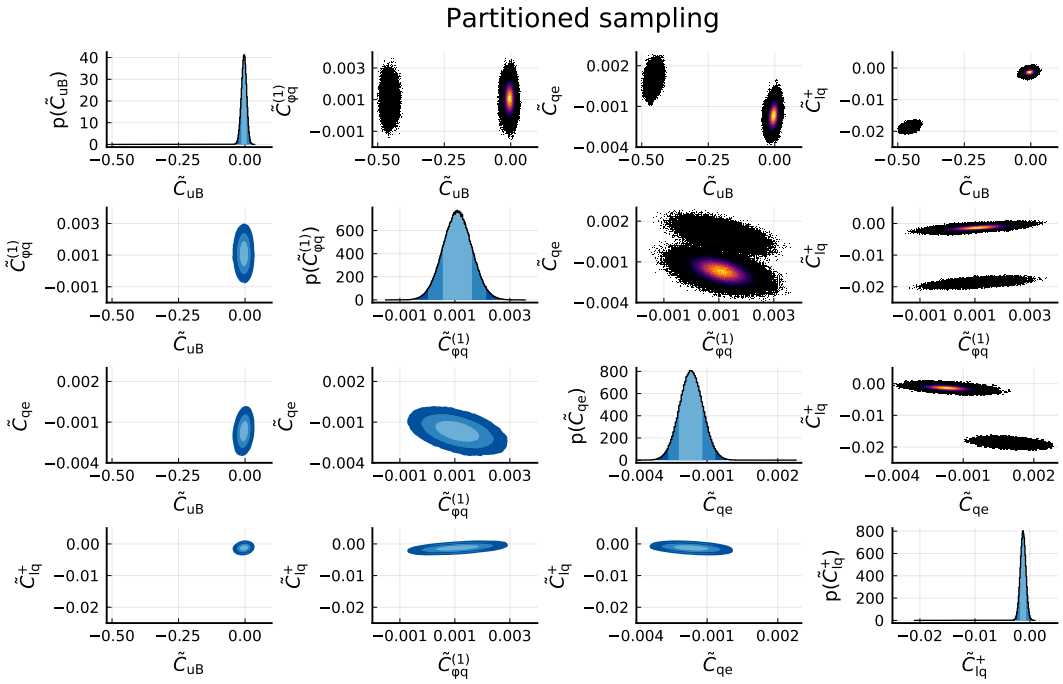

Figure E.1: Posterior distribution obtained in a fit of the four coefficients $\tilde{C}_{uB}$, $\tilde{C}^{(1)}_{\varphi q}$, $\tilde{C}_{qe}$, and $\tilde{C}^{+}_{lq}$ to the combination of current top-quark, $Z \to b\bar{b}$, and B physics data sets. One- and two-dimensional marginalized distributions are shown for posterior distributions obtained with the partitioned sampling using the Metropolis–Hastings and AHMI algorithm. The 68.3 %, 95.5 %, and 99.7 % HDRs are highlighted.

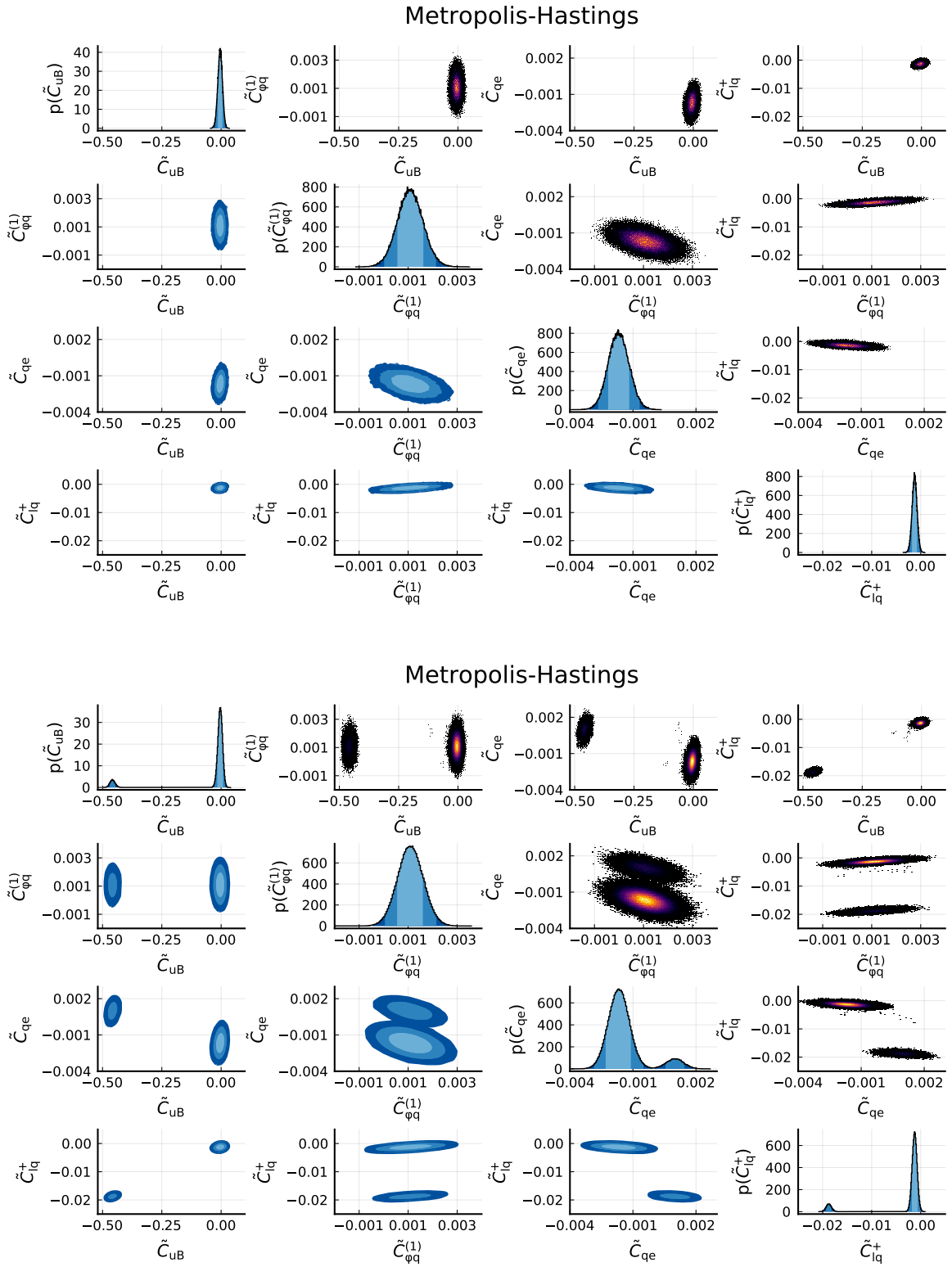

Figure E.2: Posterior distributions obtained in fits of the four coefficients $\tilde{C}_{uB}$, $\tilde{C}^{(1)}_{\varphi q}$, $\tilde{C}_{qe}$, and $\tilde{C}^{+}_{lq}$ to the combination of current top-quark, $Z \to b\bar{b}$, and B physics data sets. One- and two-dimensional marginalized distributions are shown for posterior distributions obtained in different runs of the Metropolis–Hastings algorithm. The 68.3 %, 95.5 %, and 99.7 % HDRs are highlighted.

Bibliography

[1] O. Schulz et al., *BAT.jl: A Julia-Based Tool for Bayesian Inference*, SN Computer Science **2** (2021) 210, [2008.03132].

[2] N. Castro et al., *EFTfitter: a tool for interpreting measurements in the context of effective field theories*, Eur. Phys. J. C **76** (2016) 432, [1605.05585].

[3] S. Bißmann et al., *Correlating uncertainties in global analyses within SMEFT matters*, Phys. Rev. D **102** (2020) 115019, [1912.06090].

[4] S. Bißmann et al., *Constraining top-quark couplings combining top-quark and B decay observables*, Eur. Phys. J. C **80** (2020) 136, [1909.13632].

[5] S. Bißmann et al., *Top and Beauty synergies in SMEFT-fits at present and future colliders*, JHEP **06** (2021) 010, [2012.10456].

[6] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, [1207.7214].

[7] CMS Collaboration, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, Phys. Lett. B **716** (2012) 30, [1207.7235].

[8] ATLAS Collaboration, *Search for flavour-changing neutral currents in processes with one top quark and a photon using 81 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS experiment*, Phys. Lett. B **800** (2020) 135082, [1908.08461].

[9] ATLAS Collaboration, *Measurement of the $t\bar{t}Z$ and $t\bar{t}W$ cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **99** (2019) 072009, [1901.03584].

[10] ATLAS Collaboration, *Search for flavour-changing neutral current top-quark decays $t \to qZ$ in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **07** (2018) 176, [1803.09923].

[11] CMS Collaboration, *Measurement of top quark pair production in association with a Z boson in proton-proton collisions at $\sqrt{s} = 13$ TeV*, JHEP **03** (2020) 056, [1907.11270].

[12] CMS Collaboration, *Search for the production of four top quarks in the single-lepton and opposite-sign dilepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV*, JHEP **11** (2019) 082, [1906.02805].

[13] CMS Collaboration, *Search for new physics in top quark production in dilepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **79** (2019) 886, [1903.11144].

[14] CMS Collaboration, *Measurement of the inclusive and differential $t\bar{t}+\gamma$ cross section and EFT interpretation in the single lepton channel at $\sqrt{s} = 13$ TeV*, tech. rep., CERN, 2021, URL: https://cds.cern.ch/record/2758332.

[15] D. Barducci et al., *Interpreting top-quark LHC measurements in the standard-model effective field theory*, (2018), ed. by J. A. Aguilar-Saavedra et al., [1802.07237].

[16] S. Brown et al., *TopFitter: Fitting top-quark Wilson Coefficients to Run II data*, PoS **ICHEP2018** (2019) 293, [1901.03164].

[17] I. Brivio et al., *O new physics, where art thou? A global search in the top sector*, JHEP **02** (2020) 131, [1910.03606].

[18] C. Zhang, *Effective field theory approach to top-quark decay at next-to-leading order in QCD*, Phys. Rev. D **90** (2014) 014008, [1404.1264].

[19] G. Durieux, F. Maltoni and C. Zhang, *Global approach to top-quark flavor-changing interactions*, Phys. Rev. D **91** (2015) 074017, [1412.7166].

[20] C. Degrande et al., *Single-top associated production with a Z or H boson at the LHC: the SMEFT interpretation*, JHEP **10** (2018) 005, [1804.07773].

[21] O. Bessidskaia Bylund et al., *Probing top quark neutral couplings in the Standard Model Effective Field Theory at NLO in QCD*, JHEP **05** (2016) 052, [1601.08193].

[22] F. Maltoni, L. Mantani and K. Mimasu, *Top-quark electroweak interactions at high energy*, JHEP **10** (2019) 004, [1904.05637].

[23] N. P. Hartland et al., *A Monte Carlo global analysis of the Standard Model Effective Field Theory: the top quark sector*, JHEP **04** (2019) 100, [1901.05965].

[24] A. Hald, *A history of probability and statistics and their applications before 1750*, John Wiley & Sons, 2005.

[25] A. Kolmogorov, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Ergebnisse der Mathematik und ihrer Grenzgebiete, 1933.

[26] R. T. Cox, *Probability, Frequency and Reasonable Expectation*, Am. J. Phys. **14** (1946) 1.

[27] E. T. Jaynes, *Probability Theory: The Logic of Science*, ed. by G. L. Bretthorst, Cambridge University Press, 2003.

[28] J. L. Beck, *Bayesian system identification based on probability logic*, Structural Control and Health Monitoring **17** (2010) 825.

[29] T. Bayes and R. Price, *An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S.*, Philosophical Transactions **53** (1763) 370.

[30] S. B. Mcgrayne, *The Theory That Would Not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines, and Emerged Triumphant from Two Centuries of Controversy*, Yale University Press, 2011.

[31] J. O. Berger and D. A. Berry, *Statistical Analysis and the Illusion of Objectivity*, American Scientist **76** (1988) 159, URL: http://www.jstor.org/stable/27855070.

[32] M. J. Bayarri and J. O. Berger, *The Interplay of Bayesian and Frequentist Analysis*, Statistical Science **19** (2004) 58.

[33] B. Efron, *Why Isn't Everyone a Bayesian?*, The American Statistician **40** (1986) 1, URL: http://www.jstor.org/stable/2683105.

[34] R. D. Cousins, *Why isn't every physicist a Bayesian?*, Am. J. Phys. **63** (1995) 398.

[35] G. D'Agostini, "Bayesian Reasoning Versus Conventional Statistics in High Energy Physics", *Maximum Entropy and Bayesian Methods Garching, Germany 1998*, ed. by W. von der Linden et al., Springer Netherlands, (1999) 157.

[36] L. Lyons, *Bayes and Frequentism: a Particle Physicist's perspective*, Contemp. Phys. **54** (2013) 1, [1301.1273].

[37] H. Jeffreys, *Theory of Probability*, Oxford University Press, 1939.

[38] D. V. Lindley, *A statistical paradox*, Biometrika **44** (1957) 187.

[39] H. Jeffreys, *Theory of Probability*, Oxford University Press, 1961.

[40] J. M. Bernardo, *Reference Posterior Distributions for Bayesian Inference*, Journal of the Royal Statistical Society. Series B (Methodological) **41** (1979) 113, URL: http://www.jstor.org/stable/2985028.

[41] J. Berger and J. Bernardo, *On the development of reference priors*, Bayesian statistics **4** (1992) 35.

[42] J. Berger, *The case for objective Bayesian analysis*, Bayesian analysis **1** (2006) 385.

[43] M. Betancourt, *A Conceptual Introduction to Hamiltonian Monte Carlo*, (2018), [1701.02434].

[44] B. Carpenter, *Typical Sets and the Curse of Dimensionality*, URL: https://mc-stan.org/users/documentation/case-studies/curse-dims.html, (accessed: 28.05.2021).

[45] N. Metropolis et al., *Equation of State Calculations by Fast Computing Machines*, The Journal of Chemical Physics **21** (1953) 1087.

[46] C. J. Geyer, "Introduction to Markov Chain Monte Carlo", *Handbook of Markov Chain Monte Carlo*, CRC Press, 2011, URL: https://www.routledgehandbooks.com/doi/10.1201/b10905-3.

[47] W. K. Hastings, *Monte Carlo sampling methods using Markov chains and their applications*, Biometrika **57** (1970) 97.

[48] B. Carpenter et al., *Stan: A Probabilistic Programming Language*, Journal of Statistical Software, Articles **76** (2017) 1.

[49] J. Salvatier, T. V. Wiecki and C. Fonnesbeck, *Probabilistic programming in Python using PyMC3*, PeerJ Computer Science **2** (2016) e55.

[50] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, 2017, URL: `https://www.R-project.org/`.

[51] D. J. Lunn et al., *WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility*, Statistics and Computing **10** (2000) 325.

[52] A. Caldwell, D. Kollár and K. Kröninger, *BAT: The Bayesian Analysis Toolkit*, Comput. Phys. Commun. **180** (2009) 2197.

[53] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. A **389** (1997) 81, ed. by M. Werlen and D. Perret-Gallix.

[54] A. Crivellin, M. Hoferichter and C. A. Manzari, *Fermi Constant from Muon Decay Versus Electroweak Fits and Cabibbo-Kobayashi-Maskawa Unitarity*, Phys. Rev. Lett. **127** (2021) 071801.

[55] A. Crivellin et al., *Global Electroweak Fit and Vector-Like Leptons in Light of the Cabibbo Angle Anomaly*, JHEP **12** (2020) 166, [2008.01113].

[56] O. Eberhardt, A. P. Martínez and A. Pich, *Global fits in the Aligned Two-Higgs-Doublet model*, JHEP **05** (2021) 005, [2012.09200].

[57] ATLAS Collaboration, *Search for high-mass dilepton resonances in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, Phys. Rev. D **90** (2014) 052005, [1405.4123].

[58] ATLAS Collaboration, *Search for new phenomena in the dijet mass distribution using $p - p$ collision data at $\sqrt{s} = 8$ TeV with the ATLAS detector*, Phys. Rev. D **91** (2015) 052007, [1407.1376].

[59] CUORE Collaboration, *First Results from CUORE: A Search for Lepton Number Violation via $0\nu\beta\beta$ Decay of ^{130}Te*, Phys. Rev. Lett. **120** (2018) 132501, [1710.07988].

[60] GERDA Collaboration, *Background-free search for neutrinoless double-β decay of ^{76}Ge with GERDA*, Nature **544** (2017) 47, [1703.00570].

[61] O. Luongo, G. B. Pisani and A. Troisi, *Cosmological degeneracy versus cosmography: a cosmographic dark energy model*, Int. J. Mod. Phys. D **26** (2016) 1750015, [1512.07076].

[62] P. Ullio and M. Valli, *A critical reassessment of particle Dark Matter limits from dwarf satellites*, JCAP **07** (2016) 025, [1603.07721].

[63] C. Rappold et al., *Hypernuclear production cross section in the reaction of ^{6}Li + ^{12}C at 2A GeV*, Phys. Lett. B **747** (2015) 129.

[64] J. Bezanson et al., *Julia: A fresh approach to numerical computing*, SIAM Review **59** (2017) 65.

[65] M. Besançon et al., *Distributions.jl: Definition and Modeling of Probability Distributions in the JuliaStats Ecosystem*, (2019), [1907.08611].

[66] I. Sobol, *Uniformly distributed sequences with an additional uniform property*, USSR Computational Mathematics and Mathematical Physics **16** (1976) 236.

[67] A. Gelman and D. B. Rubin, *Inference from Iterative Simulation Using Multiple Sequences*, Statistical Science **7** (1992) 457.

[68] S. P. Brooks and A. Gelman, *General Methods for Monitoring Convergence of Iterative Simulations*, Journal of Computational and Graphical Statistics **7** (1998) 434.

[69] S. Duane et al., *Hybrid Monte Carlo*, Phys. Lett. B **195** (1987) 216.

[70] R. M. Neal, "MCMC Using Hamiltonian Dynamics", *Handbook of Markov Chain Monte Carlo*, CRC Press, 2011, chap. 5.

[71] R. De Vogelaere, *Methods of integration which preserve the contact transformation property of the Hamilton equations*, tech. rep. (University of Notre Dame. Dept. of Mathematics) (1956).

[72] R. D. Ruth, *A canonical integration technique*, IEEE Trans. Nucl. Sci. **30** (1983) 2669, URL: http://cds.cern.ch/record/143981.

[73] M. Betancourt, *Identifying the Optimal Integration Time in Hamiltonian Monte Carlo*, (2016), [1601.00225].

[74] M. D. Hoffman and A. Gelman, *The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo*, Journal of Machine Learning Research **15** (2014) 1593.

[75] H. Ge, K. Xu and Z. Ghahramani, "Turing: a language for flexible probabilistic inference", *International Conference on Artificial Intelligence and Statistics, AISTATS 2018, 9-11 April 2018, Playa Blanca, Lanzarote, Canary Islands, Spain*, 2018 1682.

[76] J. Revels, M. Lubin and T. Papamarkou, *Forward-Mode Automatic Differentiation in Julia*, 2016, [1607.07892].

[77] M. Innes, *Don't Unroll Adjoint: Differentiating SSA-Form Programs*, 2019, [1810.07951].

[78] T. Hahn, *Cuba—a library for multidimensional numerical integration*, Comput. Phys. Commun. **168** (2005) 78.

[79] G. Peter Lepage, *A new algorithm for adaptive multidimensional integration*, Journal of Computational Physics **27** (1978) 192.

[80] A. Caldwell et al., *Integration with an adaptive harmonic mean algorithm*, Int. J. Mod. Phys. A **35** (2020) 1950142, [1808.08051].

[81] V. Hafych et al., *Parallelizing MCMC Sampling via Space Partitioning*, 2020, [2008.03098].

[82] J. Skilling, "Nested Sampling", *Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 24th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, ed. by R. Fischer, R. Preuss and U. V. Toussaint, vol. 735, American Institute of Physics Conference Series, 2004 395.

[83] J. Buchner, *Nested Sampling Methods*, 2021, [2101.09675].

[84] M. Betancourt, "Nested sampling with constrained hamiltonian monte carlo", *AIP Conference Proceedings*, vol. 1305, 1, American Institute of Physics, (2011) 165.

[85] J. Buchner, *A statistical test for Nested Sampling algorithms*, Statistics and Computing **26** (2014) 383.

[86] J. Buchner, *Collaborative Nested Sampling: Big Data versus Complex Physical Models*, Publications of the Astronomical Society of the Pacific **131** (2019) 108005.

[87] M. Lucas et al., *TuringLang/NestedSamplers.jl: v0.6.3*, version v0.6.3, 2021, URL: https://doi.org/10.5281/zenodo.4718970.

[88] P. K. Mogensen and A. N. Riseth, *Optim: A mathematical optimization package for Julia*, Journal of Open Source Software **3** (2018) 615.

[89] J. Nelder and R. Mead, *A Simplex Method for Function Minimization*, Comput. J. **7** (1965) 308.

[90] D. Liu and J. Nocedal, *On the limited memory BFGS method for large scale optimization*, Mathematical Programming **45** (1989) 503.

[91] D. Freedman and P. Diaconis, *On the histogram as a density estimator: L 2 theory*, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete **57** (1981) 453.

[92] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579.

[93] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[94] S. Weinberg, *Nonabelian Gauge Theories of the Strong Interactions*, Phys. Rev. Lett. **31** (1973) 494.

[95] H. Fritzsch, M. Gell-Mann and H. Leutwyler, *Advantages of the Color Octet Gluon Picture*, Phys. Lett. B **47** (1973) 365.

[96] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. C **680519** (1968) 367.

[97] G. 't Hooft and M. J. G. Veltman, *Regularization and Renormalization of Gauge Fields*, Nucl. Phys. B **44** (1972) 189.

[98] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132.

[99] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508, ed. by J. C. Taylor.

[100] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321, ed. by J. C. Taylor.

[101] Particle Data Group, *Review of Particle Physics*, PTEP **2020** (2020) 083C01.

[102] Super-Kamiokande Collaboration, *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562, [hep-ex/9807003].

[103] SNO Collaboration, *Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **89** (2002) 011301, [nucl-ex/0204008].

[104] KamLAND Collaboration, *First results from KamLAND: Evidence for reactor anti-neutrino disappearance*, Phys. Rev. Lett. **90** (2003) 021802, [hep-ex/0212021].

[105] K2K Collaboration, *Measurement of Neutrino Oscillation by the K2K Experiment*, Phys. Rev. D **74** (2006) 072003, [hep-ex/0606032].

[106] B. Pontecorvo, *Neutrino Experiments and the Problem of Conservation of Leptonic Charge*, Zh. Eksp. Teor. Fiz. **53** (1967) 1717.

[107] Z. Maki, M. Nakagawa and S. Sakata, *Remarks on the Unified Model of Elementary Particles*, Progress of Theoretical Physics **28** (1962) 870.

[108] G. Bertone, D. Hooper and J. Silk, *Particle dark matter: Evidence, candidates and constraints*, Phys. Rept. **405** (2005) 279, [hep-ph/0404175].

[109] D. Clowe et al., *A direct empirical proof of the existence of dark matter*, Astrophys. J. Lett. **648** (2006) L109, [astro-ph/0608407].

[110] ESA and the Planck Collaboration, *Planck's new cosmic recipe*, URL: https://sci.esa.int/web/planck/-/51557-planck-new-cosmic-recipe, (accessed: 25.05.2021).

[111] BaBar Collaboration, *Evidence for an excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays*, Phys. Rev. Lett. **109** (2012) 101802, [1205.5442].

[112] BaBar Collaboration, *Measurement of an Excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ Decays and Implications for Charged Higgs Bosons*, Phys. Rev. D **88** (2013) 072012, [1303.0571].

[113] Belle Collaboration, *Measurement of the branching ratio of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ relative to $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays with hadronic tagging at Belle*, Phys. Rev. D **92** (2015) 072014, [1507.03233].

[114] Belle Collaboration, "Measurement of the branching ratio of $\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau$ relative to $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$ decays with a semileptonic tagging method", *51st Rencontres de Moriond on EW Interactions and Unified Theories*, 2016, [1603.06711].

[115] LHCb Collaboration, *Test of lepton universality in beauty-quark decays*, (2021), [2103.11769].

[116] LHCb Collaboration, *Branching fraction measurements of the rare $B_s^0 \to \phi\mu^+\mu^-$ and $B_s^0 \to f_2'(1525)\mu^+\mu^-$ decays*, (2021), [2105.14007].

[117] Muon g-2 Collaboration, *Measurement of the negative muon anomalous magnetic moment to 0.7 ppm*, Phys. Rev. Lett. **92** (2004) 161802, [hep-ex/0401008].

[118] Muon g-2 Collaboration, *Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm*, Phys. Rev. Lett. **126** (2021) 141801, [2104.03281].

[119] S. Borsanyi et al., *Leading hadronic contribution to the muon 2 magnetic moment from lattice QCD*, Nature **593** (2021) 51, [2002.12347].

[120] C. Degrande, *Effective field theories in the standard model and beyond*, PhD thesis, Louvain U. - CP3, 2011.

[121] J. Brehmer, *New Ideas for Effective Higgs Measurements*, PhD thesis, U. Heidelberg, 2017.

[122] E. Fermi, *Versuch einer Theorie der β-Strahlen. I*, Zeitschrift für Physik **88** (1934) 161.

[123] UA1 Collaboration, *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540\,GeV$*, Phys. Lett. B **122** (1983) 103.

[124] UA2 Collaboration, *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider*, Phys. Lett. B **122** (1983) 476.

[125] A. J. Buras et al., *Theoretical uncertainties and phenomenological aspects of $B \to X_s\gamma$ decay*, Nucl. Phys. B **424** (1994) 374, [hep-ph/9311345].

[126] K. G. Chetyrkin, M. Misiak and M. Munz, *Weak radiative B meson decay beyond leading logarithms*, Phys. Lett. B **400** (1997) 206, [Erratum: Phys. Lett. B 425, 414 (1998)], [hep-ph/9612313].

[127] J. Aebischer et al., *B physics Beyond the Standard Model at One Loop: Complete Renormalization Group Evolution below the Electroweak Scale*, JHEP **09** (2017) 158, [1704.06639].

[128] E. E. Jenkins, A. V. Manohar and P. Stoffer, *Low-Energy Effective Field Theory below the Electroweak Scale: Operators and Matching*, JHEP **03** (2018) 016, [1709.04486].

[129] I. Brivio and M. Trott, *The Standard Model as an Effective Field Theory*, Phys. Rept. **793** (2019) 1, [1706.08945].

[130] W. Buchmuller and D. Wyler, *Effective Lagrangian Analysis of New Interactions and Flavor Conservation*, Nucl. Phys. B **268** (1986) 621.

[131] C. Degrande et al., *Effective Field Theory: A Modern Approach to Anomalous Couplings*, Annals Phys. **335** (2013) 21, [1205.4231].

[132] A. Kobach, *Baryon Number, Lepton Number, and Operator Dimension in the Standard Model*, Phys. Lett. B **758** (2016) 455, [1604.05726].

[133] B. Grzadkowski et al., *Dimension-Six Terms in the Standard Model Lagrangian*, JHEP **10** (2010) 085, [1008.4884].

[134] S. Fichet, A. Tonero and P. Rebello Teles, *Sharpening the shape analysis for higher-dimensional operator searches*, Phys. Rev. D **96** (2017) 036003, [1611.01165].

[135] M. Cristinziani and M. Mulders, *Top-quark physics at the Large Hadron Collider*, J. Phys. G **44** (2017) 063001, [1606.00327].

[136] A. Buckley et al., *Constraining top quark effective theory in the LHC Run II era*, JHEP **04** (2016) 015, [1512.03360].

[137] S. Brown et al., *New results from TopFitter*, PoS **ICHEP2020** (2021) 322.

[138] ATLAS and CMS Collaborations, *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV*, JHEP **08** (2016) 045, [1606.02266].

[139] ATLAS and CMS Collaborations, *Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments*, Phys. Rev. Lett. **114** (2015) 191803, [1503.07589].

[140] A. Buckley et al., *Global fit of top quark effective theory to data*, Phys. Rev. D **92** (2015) 091501, [1506.08845].

[141] L. Lyons, D. Gibaut and P. Clifford, *How to Combine Correlated Estimates of a Single Physical Quantity*, Nucl. Instrum. Meth. A **270** (1988) 110.

[142] A. Valassi, *Combining correlated measurements of several different physical quantities*, Nucl. Instrum. Meth. A **500** (2003) 391.

[143] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, Annals Math. Statist. **9** (1938) 60.

[144] A. Valassi and R. Chierici, *Information and treatment of unknown correlations in the combination of measurements using the BLUE method*, Eur. Phys. J. C **74** (2014).

[145] CDF and D0 Collaborations and Tevatron Electroweak Working Group, *Combination of CDF and D0 results on the mass of the top quark using up to 5.8 fb^{-1} of data*, (2011), [1107.5255].

[146] ATLAS Collaboration, *Combination of ATLAS and CMS results on the mass of the top quark using up to 4.9 fb^{-1} of data*, tech. rep. ATLAS-CONF-2012-095, CERN, 2012, URL: https://cds.cern.ch/record/1460441.

[147] R. J. Hyndman, *Computing and Graphing Highest Density Regions*, The American Statistician **50** (1996) 120, URL: http://www.jstor.org/stable/2684423.

[148] ATLAS Collaboration, *Comprehensive measurements of t-channel single top-quark production cross sections at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Rev. D **90** (2014) 112006.

[149] CMS Collaboration, *Measurement of the Single-Top-Quark t-Channel Cross Section in pp Collisions at $\sqrt{s} = 7$ TeV*, JHEP **12** (2012) 035, [1209.4533].

[150] ATLAS Collaboration, *Fiducial, total and differential cross-section measurements of t-channel single top-quark production in pp collisions at 8 TeV using data collected by the ATLAS detector*, Eur. Phys. J. C **77** (2017) 531, [1702.02859].

[151] CMS Collaboration, *Measurement of the t-channel single-top-quark production cross section and of the $|V_{tb}|$ CKM matrix element in pp collisions at $\sqrt{s} = 8$ TeV*, JHEP **06** (2014) 090, [1403.7366].

[152] CMS Collaboration, *Single top t-channel differential cross section at 8 TeV*, tech. rep. CMS-PAS-TOP-14-004, CERN, 2014, URL: https://cds.cern.ch/record/1956681.

[153] ATLAS Collaboration, *Measurement of the inclusive cross-sections of single top-quark and top-antiquark t-channel production in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **04** (2017) 086, [1609.03920].

[154] CMS Collaboration, *Cross section measurement of t-channel single top quark production in pp collisions at $\sqrt{s} = 13$ TeV*, Phys. Lett. B **772** (2017) 752, [1610.00678].

[155] CMS Collaboration, *Measurement of the differential cross section for t-channel single-top-quark production at $\sqrt{s} = 13$ TeV*, tech. rep. CMS-PAS-TOP-16-004, CERN, 2016, URL: https://cds.cern.ch/record/2151074.

[156] CMS Collaboration, *Measurement of the single top quark and antiquark production cross sections in the t channel and their ratio in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Phys. Lett. B **800** (2020) 135042, [1812.10514].

[157] CDF Collaboration, *Measurement of W-Boson Polarization in Top-quark Decay in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV*, Phys. Rev. Lett. **105** (2010) 042002, [1003.0224].

[158] CDF Collaboration, *Measurement of W-Boson Polarization in Top-quark Decay using the Full CDF Run II Data Set*, Phys. Rev. D **87** (2013) 031104, [1211.4523].

[159] D0 Collaboration, *Measurement of the W Boson Helicity in Top Quark Decays using 5.4 fb^{-1} of $p\bar{p}$ Collision Data*, Phys. Rev. D **83** (2011) 032009, [1011.6549].

[160] ATLAS Collaboration, *Measurement of the W boson polarization in top quark decays with the ATLAS detector*, JHEP **06** (2012) 088, [1205.2484].

[161] CMS Collaboration, *Measurement of the W-Boson Helicity in Top-Quark decays from $t\bar{t}$ Production in Lepton + Jets Events in pp Collisions at $\sqrt{s} = 7$ TeV*, JHEP **10** (2013) 167, [1308.3879].

[162] ATLAS Collaboration, *Direct top-quark decay width measurement in the $t\bar{t}$ lepton+jets channel at $\sqrt{s} = 8$ TeV with the ATLAS experiment*, Eur. Phys. J. C **78** (2018) 129, [1709.04207].

[163] ATLAS Collaboration, *Measurement of the W boson polarisation in $t\bar{t}$ events from pp collisions at $\sqrt{s} = 8$ TeV in the lepton + jets channel with ATLAS*, Eur. Phys. J. C **77** (2017) 264, [Erratum: Eur. Phys. J. C 79, 19 (2019)], [1612.02577].

[164] CMS Collaboration, *Measurement of the W boson helicity in events with a single reconstructed top quark in pp collisions at $\sqrt{s} = 8$ TeV*, JHEP **01** (2015) 053, [1410.1154].

[165] CMS Collaboration, *Measurement of the W boson helicity fractions in the decays of top quark pairs to lepton + jets final states produced in pp collisions at $\sqrt{s} = 8$ TeV*, Phys. Lett. B **762** (2016) 512, [1605.09047].

[166] D0 Collaboration, *Observation of Single Top Quark Production*, Phys. Rev. Lett. **103** (2009) 092001, [0903.0850].

[167] CDF Collaboration, *First Observation of Electroweak Single Top Quark Production*, Phys. Rev. Lett. **103** (2009) 092002, [0903.0885].

[168] T. M. P. Tait and C. .-P. Yuan, *Single top quark production as a window to physics beyond the standard model*, Phys. Rev. D **63** (2000) 014018, [hep-ph/0007298].

[169] M. de Beurs et al., *Effective operators in t-channel single top production and decay*, Eur. Phys. J. C **78** (2018) 919, [1807.03576].

[170] J. Ellis et al., *Top, Higgs, Diboson and Electroweak Fit to the Standard Model Effective Field Theory*, JHEP **04** (2021) 279, [2012.02779].

[171] J. J. Ethier et al., *Combined SMEFT interpretation of Higgs, diboson, and top quark data from the LHC*, (2021), [2105.00006].

[172] J. M. Campbell et al., *Next-to-Leading-Order Predictions for t-Channel Single-Top Production at Hadron Colliders*, Phys. Rev. Lett. **102** (2009).

[173] D0 Collaboration, *An Improved determination of the width of the top quark*, Phys. Rev. D **85** (2012) 091104, [1201.4156].

[174] CMS Collaboration, *Measurement of the ratio $\mathcal{B}(t \to Wb)/\mathcal{B}(t \to Wq)$ in pp collisions at $\sqrt{s} = 8$ TeV*, Phys. Lett. B **736** (2014) 33, [1404.2292].

[175] ATLAS Collaboration, *Measurement of the top-quark decay width in top-quark pair events in the dilepton channel at $\sqrt{s} = 13$ TeV with the ATLAS detector*, tech. rep., CERN, 2019, URL: http://cds.cern.ch/record/2684952.

[176] G. L. Kane, G. A. Ladinsky and C. .-P. Yuan, *Using the top quark for testing standard-model polarization and* CP *predictions*, Phys. Rev. D **45** (1992) 124.

[177] A. Czarnecki, J. G. Körner and J. H. Piclum, *Helicity fractions of W bosons from top quark decays at next-to-next-to-leading order in QCD*, Phys. Rev. D **81** (2010) 111503, [1005.2625].

[178] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, [1405.0301].

[179] C. Degrande et al., *UFO - The Universal FeynRules Output*, Comput. Phys. Commun. **183** (2012) 1201, [1108.2040].

[180] A. D. Martin et al., *Parton distributions for the LHC*, Eur. Phys. J. C **63** (2009) 189, [0901.0002].

[181] H.-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. D **82** (2010) 074024, [1007.2241].

[182] NNPDF Collaboration, *Parton distributions with QED corrections*, Nucl. Phys. B **877** (2013) 290, [1308.0598].

[183] C. Zhang and S. Willenbrock, *Effective-Field-Theory Approach to Top-Quark Production and Decay*, Phys. Rev. D **83** (2011) 034006, [1008.3869].

[184] J. Gao, C. S. Li and H. X. Zhu, *Top Quark Decay at Next-to-Next-to Leading Order in QCD*, Phys. Rev. Lett. **110** (2013) 042001, [1210.2808].

[185] ATLAS and CMS Collaborations, *Combinations of single-top-quark production cross-section measurements and $|f_{LV}V_{tb}|$ determinations at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS experiments*, JHEP **05** (2019) 088, [1902.07158].

[186] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, tech. rep., CERN, 2019, URL: http://cds.cern.ch/record/2677054.

[187] A. G. et al., *High-Luminosity Large Hadron Collider (HL-LHC): Technical Design Report V. 0.1*, CERN Yellow Reports: Monographs, CERN, 2017, URL: https://cds.cern.ch/record/2284929.

[188] R. Aoude et al., *The impact of flavour data on global fits of the MFV SMEFT*, JHEP **12** (2020) 113, [2003.05432].

[189] P. J. Fox et al., *Deciphering top flavor violation at the LHC with B factories*, Phys. Rev. D **78** (2008) 054008, [0704.1482].

[190] B. Grzadkowski and M. Misiak, *Anomalous Wtb coupling effects in the weak radiative B-meson decay*, Phys. Rev. D **78** (2008) 077501, [Erratum: Phys. Rev. D 84, 059903 (2011)], [0802.1413].

[191] J. Drobnak, S. Fajfer and J. F. Kamenik, *Probing anomalous tWb interactions with rare B decays*, Nucl. Phys. B **855** (2012) 82, [1109.2357].

[192] J. Brod et al., *Probing anomalous $t\bar{t}Z$ interactions with rare meson decays*, JHEP **02** (2015) 141, [1408.0792].

[193] S. Moch and P. Uwer, *Theoretical status and prospects for top-quark pair production at hadron colliders*, Phys. Rev. D **78** (2008) 034003, [0804.1476].

[194] J. Aebischer et al., *Matching of gauge invariant dimension-six operators for $b \to s$ and $b \to c$ transitions*, JHEP **2016** (2016).

[195] E. E. Jenkins, A. V. Manohar and M. Trott, *Naive Dimensional Analysis Counting of Gauge Theory Amplitudes and Anomalous Dimensions*, Phys. Lett. B **726** (2013) 697, [1309.0819].

[196] E. E. Jenkins, A. V. Manohar and M. Trott, *Renormalization Group Evolution of the Standard Model Dimension Six Operators I: Formalism and lambda Dependence*, JHEP **10** (2013) 087, [1308.2627].

[197] E. E. Jenkins, A. V. Manohar and M. Trott, *Renormalization Group Evolution of the Standard Model Dimension Six Operators II: Yukawa Dependence*, JHEP **01** (2014) 035, [1310.4838].

[198] R. Alonso et al., *Renormalization Group Evolution of the Standard Model Dimension Six Operators III: Gauge Coupling Dependence and Phenomenology*, JHEP **04** (2014) 159, [1312.2014].

[199] J. Aebischer, J. Kumar and D. M. Straub, *Wilson: a Python package for the running and matching of Wilson coefficients above and below the electroweak scale*, Eur. Phys. J. C **78** (2018) 1026, [1804.05033].

[200] D. M. Straub, *flavio: a Python package for flavour and precision phenomenology in the Standard Model and beyond*, (2018), [1810.08132].

[201] ATLAS Collaboration, *Measurements of inclusive and differential fiducial cross-sections of $t\bar{t}\gamma$ production in leptonic final states at $\sqrt{s} = 13$ TeV in ATLAS*, Eur. Phys. J. C **79** (2019) 382, [1812.01697].

[202] ATLAS Collaboration, *Measurement of the $t\bar{t}\gamma$ production cross section in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, JHEP **11** (2017) 086, [1706.03046].

[203] ATLAS Collaboration, *Observation of top-quark pair production in association with a photon and measurement of the $t\bar{t}\gamma$ production cross section in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, Phys. Rev. D **91** (2015) 072007, [1502.00586].

[204] CMS Collaboration, *Measurement of the semileptonic $t\bar{t}$ + γ production cross section in pp collisions at $\sqrt{s} = 8$ TeV*, JHEP **10** (2017) 006, [1706.08128].

[205] CMS Collaboration, *Measurement of the inclusive and differential $t\bar{t} + \gamma$ cross section and EFT interpretation in the single lepton channel at $\sqrt{s} = 13$ TeV*, tech. rep., 2021, URL: https://cds.cern.ch/record/2758332.

[206] K. Melnikov, M. Schulze and A. Scharf, *QCD corrections to top quark pair production in association with a photon at hadron colliders*, Phys. Rev. D **83** (2011) 074013, [1102.1967].

[207] HFLAV Collaboration, 2019, URL: https://hflav-eos.web.cern.ch/hflav-eos/rare/April2019/RADLL/OUTPUT/HTML/radll_table3.html.

[208] M. Misiak et al., *Updated NNLO QCD predictions for the weak radiative B-meson decays*, Phys. Rev. Lett. **114** (2015) 221801, [1503.01789].

[209] BaBar Collaboration, *Measurement of the $B \to X_s\gamma$ branching fraction and photon energy spectrum using the recoil method*, Phys. Rev. D **77** (2008) 051103, [0711.4889].

[210] BaBar Collaboration, *Precision Measurement of the $B \to X_s\gamma$ Photon Energy Spectrum, Branching Fraction, and Direct CP Asymmetry $A_{CP}(B \to X_{s+d}\gamma)$*, Phys. Rev. Lett. **109** (2012) 191801, [1207.2690].

[211] BaBar Collaboration, *Exclusive Measurements of $b \to s\gamma$ Transition Rate and Photon Energy Spectrum*, Phys. Rev. D **86** (2012) 052012, [1207.2520].

[212] Belle Collaboration, *Measurement of Inclusive Radiative B-meson Decays with a Photon Energy Threshold of 1.7 GeV*, Phys. Rev. Lett. **103** (2009) 241801, [0907.1384].

[213] Belle Collaboration, *Measurement of the $\bar{B} \to X_s\gamma$ Branching Fraction with a Sum of Exclusive Decays*, Phys. Rev. D **91** (2015) 052004, [1411.7198].

[214] Belle Collaboration, "Measurement of the inclusive $B \to X_{s+d}\gamma$ branching fraction, photon energy spectrum and HQE parameters", *38th International Conference on High Energy Physics*, 2016, [1608.02344].

[215] CLEO Collaboration, *Branching fraction and photon energy spectrum for $b \to s\gamma$*, Phys. Rev. Lett. **87** (2001) 251807, [hep-ex/0108032].

[216] HFLAV Collaboration, *Averages of b-hadron, c-hadron, and τ-lepton properties as of 2018*, Eur. Phys. J. C **81** (2021) 226, [1909.12524].

[217] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, [1410.3012].

[218] E. Conte, B. Fuks and G. Serret, *MadAnalysis 5, A User-Friendly Framework for Collider Phenomenology*, Comput. Phys. Commun. **184** (2013) 222, [1206.1599].

[219] E. Conte et al., *Designing and recasting LHC analyses with MadAnalysis 5*, Eur. Phys. J. C **74** (2014) 3103, [1405.3982].

[220] B. Dumont et al., *Toward a public analysis database for LHC new physics searches using MADANALYSIS 5*, Eur. Phys. J. C **75** (2015) 56, [1407.3278].

[221] E. Conte and B. Fuks, *Confronting new physics theories to LHC data with MADANALYSIS 5*, Int. J. Mod. Phys. A **33** (2018) 1830027, [1808.00480].

[222] M. Cacciari, G. P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, [0802.1189].

[223] M. Cacciari, G. P. Salam and G. Soyez, *FastJet User Manual*, Eur. Phys. J. C **72** (2012) 1896, [1111.6097].

[224] A. Ali et al., *A Comparative study of the decays $B \to (K, K^{*)}\ell^+\ell^-$ in standard model and supersymmetric theories*, Phys. Rev. D **61** (2000) 074024, [hep-ph/9910221].

[225] LHCb Collaboration, *Differential branching fraction and angular analysis of the decay $B^0 \to K^{*0}\mu^+\mu^-$*, JHEP **08** (2013) 131, [1304.6325].

[226] L. Di Luzio et al., *ΔM_s theory precision confronts flavour anomalies*, JHEP **12** (2019) 009, [1909.11087].

[227] G. Durieux et al., *The electro-weak couplings of the top and bottom quarks — Global fit and future prospects*, JHEP **12** (2019) 98, [Erratum: JHEP 01, 195 (2021)], [1907.10619].

[228] ATLAS Collaboration, *Measurement of the $t\bar{t}$ production cross-section and lepton differential distributions in $e\mu$ dilepton events from pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. C **80** (2020) 528, [1910.08819].

[229] S. Frixione et al., *Electroweak and QCD corrections to top-pair hadroproduction in association with heavy bosons*, JHEP **06** (2015) 184, [1504.03446].

[230] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, **2/2017** (2016), [1610.07922].

[231] R. Frederix et al., *The automation of next-to-leading order electroweak calculations*, JHEP **07** (2018) 185, [1804.10017].

[232] M. Czakon and A. Mitov, *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, Comput. Phys. Commun. **185** (2014) 2930, [1112.5675].

[233] BaBar Collaboration, *Measurement of the $B \to X_s l^+ l^-$ branching fraction and search for direct CP violation from a sum of exclusive final states*, Phys. Rev. Lett. **112** (2014) 211802, [1312.5364].

[234] T. Huber, T. Hurth and E. Lunghi, *Inclusive $\bar{B} \to X_s\ell^+\ell^-$: complete angular analysis and a thorough study of collinear photons*, JHEP **06** (2015) 176, [1503.04849].

[235] Belle Collaboration, *Measurement of the lepton forward-backward asymmetry in $B \to X_s\ell^+\ell^-$ decays with a sum of exclusive modes*, Phys. Rev. D **93** (2016) 032008, [Addendum: Phys.Rev.D 93, 059901 (2016)], [1402.7134].

[236] F. Archilli, *Rare penguin decays at LHCb*, (2021), URL: https://cds.cern.ch/record/2759107.

[237] LHCb Collaboration, *Measurement of CP-Averaged Observables in the $B^0 \to K^{*0}\mu^+\mu^-$ Decay*, Phys. Rev. Lett. **125** (2020) 011802, [2003.04831].

[238] LHCb Collaboration, *Differential branching fractions and isospin asymmetries of $B \to K^{(*)}\mu^+\mu^-$ decays*, JHEP **06** (2014) 133, [1403.8044].

[239] LHCb Collaboration, *Angular analysis and differential branching fraction of the decay $B_s^0 \to \phi\mu^+\mu^-$*, JHEP **09** (2015) 179, [1506.08777].

[240] LHCb Collaboration, *Differential branching fraction and angular analysis of $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decays*, JHEP **06** (2015) 115, [Erratum: JHEP 09, 145 (2018)], [1503.07138].

[241] J. Aebischer et al., *A Global Likelihood for Precision Constraints and Flavour Anomalies*, Eur. Phys. J. C **79** (2019) 509, [1810.07698].

[242] ALEPH, DELPHI, L3, OPAL, SLD, LEP Electroweak Working Group, SLD Electroweak Group, SLD Heavy Flavour Group, *Precision electroweak measurements on the Z resonance*, Phys. Rept. **427** (2006) 257, [hep-ex/0509008].

[243] ATLAS and CMS Collaborations, *Addendum to the report on the physics at the HL-LHC, and perspectives for the HE-LHC: Collection of notes from ATLAS and CMS*, CERN Yellow Rep. Monogr. **7** (2019) Addendum, [1902.10229].

[244] Belle-II Collaboration, *The Belle II Physics Book*, PTEP **2019** (2019) 123C01, ed. by E. Kou and P. Urquijo, [Erratum: PTEP 2020, 029201 (2020)], [1808.10567].

[245] Belle-II Collaboration, *Search for $B^+ \to K^+\nu\bar{\nu}$ decays using an inclusive tagging method at Belle II*, (2021), [2104.12624].

[246] FCC Collaboration, *FCC-ee: The Lepton Collider: Future Circular Collider Conceptual Design Report Volume 2*, Eur. Phys. J. ST **228** (2019) 261.

[247] *The International Linear Collider Technical Design Report - Volume 1: Executive Summary*, (2013), ed. by T. Behnke et al., [1306.6327].

[248] CLICdp Collaboration, *Top-Quark Physics at the CLIC Electron-Positron Linear Collider*, JHEP **11** (2019) 003, [1807.02441].

[249] ATLAS Collaboration, *Prospects for the measurement of $t\bar{t}\gamma$ with the upgraded ATLAS detector at the High-Luminosity LHC*, (2018), ATL-PHYS-PUB-2018-049.

[250] CMS Collaboration, *Anomalous couplings in the tt+Z final state at the HL-LHC*, (2018), CMS-PAS-FTR-18-036.

[251] CMS Collaboration, *Projection of measurements of differential ttbar production cross sections in the e/μ+jets channels in pp collisions at the HL-LHC*, (2018), CMS-PAS-FTR-18-015.

[252] ATLAS Collaboration, *Measurements of inclusive and differential fiducial cross-sections of $t\bar{t}\gamma$ production in leptonic final states at $\sqrt{s} = 13$ TeV in ATLAS*, Eur. Phys. J. C **79** (2019) 382, [1812.01697].

[253] A. J. Buras et al., *$B \to K^{(*)}\nu\overline{\nu}$ decays in the Standard Model and beyond*, JHEP **02** (2015) 184, [1409.4557].

[254] A. M. Sirunyan et al., *Search for new physics in top quark production with additional leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV using effective field theory*, JHEP **2021** (2021).

[255] G. Durieux et al., *Global and optimal probes for the top-quark effective field theory at future lepton colliders*, JHEP **10** (2018) 168, [1807.02121].

[256] H. Tak, X.-L. Meng and D. A. van Dyk, *A Repelling–Attracting Metropolis Algorithm for Multimodality*, Journal of Computational and Graphical Statistics **27** (2018) 479.

[257] E. Pompe, C. Holmes and K. Łatuszyński, *A framework for adaptive MCMC targeting multimodal distributions*, The Annals of Statistics **48** (2020) 2930.

[258] H. Tjelmeland and B. K. Hegstad, *Mode Jumping Proposals in MCMC*, Scandinavian Journal of Statistics **28** (2001) 205, URL: http://www.jstor.org/stable/4616652.

List of Abbreviations

4FS four-flavor scheme

5FS five-flavor scheme

AHMI Adaptive Harmonic Mean Integration

BAT Bayesian Analysis Toolkit

BAT.jl Bayesian Analysis Toolkit in Julia

BLUE Best Linear Unbiased Estimator

BSM physics beyond the Standard Model

CLIC Compact Linear Collider

CVWs central value weights

EFT effective field theory

ESS effective sample size

FCC Future Circular Collider

HDA highest density area

HDI highest density interval

HDL highest density length

HDR highest density region

HDV highest density volume

HEFT Higgs effective field theory

HFLAV Heavy Flavor Averaging Group

HL-LHC High-Luminosity LHC

HMC Hamiltonian Monte Carlo

iid independent and identically distributed

IIW intrinsic information weight

ILC International Linear Collider

LEFT low-energy effective field theory

LHC Large Hadron Collider

LO leading order

LRPS likelihood-restricted prior sampling

MC Monte Carlo

MCMC Markov chain Monte Carlo

MEW measurement exclusion weight

MIW marginal information weight

NLO next-to-leading order

NNLO next-to-next-to-leading order

NUTS No-U-Turn sampler

PDF parton distribution function

PSRF potential scale reduction factor

QCD quantum chromodynamics

QED quantum electrodynamics

RGE renormalization group equation

RIW relative importance weight

SM Standard Model of particle physics

SMEFT Standard Model effective field theory

SSB spontaneous symmetry breaking

UEW uncertainty exclusion weight

UFO Universal FeynRules Output

UV ultraviolet

WET weak effective theory

Acknowledgments

This thesis would not have been possible without the help and support of so many people, whom I want to thank at this point.

First of all, I would like to thank Prof. Dr. Kevin Kröninger for giving me the opportunity of writing my thesis in his working group. He has been a great supervisor and always found time to discuss some physics, despite all his administrative duties. He supported me in so many ways, and I especially want to thank him for giving me the opportunities to attend several conferences and schools in nice places around the world. I would like to thank Prof. Dr. Allen Caldwell for agreeing to be the second referee of this thesis and for the excellent collaboration with him and his team on the BAT.jl project. I also want to thank Prof. Dr. Gudrun Hiller for the very successful cooperation and all her expert advice on our EFT projects. I would like to acknowledge the support by the *Studienstiftung des deutschen Volkes*. The scholarship did not only enable the funding of this thesis, but also provided many inspiring experiences at various events.

Throughout my time at E4, I received a lot of competent advice from Priv.-Doz. Dr. Johannes Erdmann, regarding various projects and topics. I want to thank him for all of this and for the very good cooperation on our projects. I would like to thank Dr. Oliver Schulz for the very enjoyable collaboration on BAT.jl and all his valuable feedback, in particular, for the technical advice on programming in Julia, which was a great help for me. I also want to thank all members of the BAT.jl team for the very good cooperation over the last years and some nice evenings we enjoyed together in Munich and Dortmund.

Many parts of the work presented in this thesis would not have been possible without Dr. Stefan Bißmann. I owe him a very big *thank you* for the great collaboration over the last few years. It has been a pleasure to work so closely with him and talk to him on a daily basis about all the crazy EFT stuff, but of course also about a lot of other things beyond physics. I am very grateful that he had the composure to explain certain theoretical details to me again and again when I asked him the same questions over and over. Of course, it is needless to say that I very much enjoyed our collaboration and, in particular, the schools we visited together, even the ones during an emerging pandemic. I would also like to thank Salvatore and Lars for the good cooperation and often amusing virtual meetings that broke through the monotony of the home office during the last year.

In particular, I want to thank Johannes, Oliver, Olaf, Stefan, Salvatore, Lars, Jan Lukas, Björn, and Kevin for proofreading parts of this thesis and providing very helpful feedback.

A lot of nice people made working, and in particular not-working, at E4 very enjoyable. I would especially like to thank all the former residents of *"Office-137"* for the great times we had in our office. I also want to thank everyone who made the after-work and *"E4 e.V."* social events so much fun. Special thanks also to Andrea, Mike, and Markus, who kept the things running at E4 during my time here.

I want to thank my friends of the *"Freundestorus"* for the great time we had studying physics in Dortmund together. I am really glad that we met nine years ago and that our group still manages to meet a couple of times a year, even though people are now spread across Germany (and Switzerland).

I am infinitely grateful to my parents and my family for all their support and help I received during the years of my PhD and throughout my whole life. Of course, the final words of acknowledgments are dedicated to my beloved wife (and inevitable home office colleague) Sandra. Without you, all your support and love, I would not have been able to finish this thesis, and my life would definitely not be as good as it is.