

Quantum fitting framework applied to effective field theories

Juan Carlos Criado^{*} and Michael Spannowsky[†]

*Institute for Particle Physics Phenomenology, Durham University, Durham DH1 3LE, United Kingdom
and Department of Physics, Durham University, Durham DH1 3LE, United Kingdom*

Roman Kogler[‡]

Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, 22607 Hamburg, Germany



(Received 21 September 2022; accepted 5 January 2023; published 25 January 2023)

The use of experimental data to constrain the values of the Wilson coefficients of an effective field theory involves minimizing a χ^2 function that may contain local minima. Classical optimization algorithms can become trapped in these minima, preventing the determination of the global minimum. The quantum annealing framework has the potential to overcome this limitation and reliably find the global minimum of nonconvex functions. We present QFitter, a quantum annealing method to perform effective field theory fits. Using a state-of-the-art quantum annealer, we show with concrete examples that QFitter can be used to fit sets of at least eight coefficients, including their quadratic contributions. An arbitrary number of observables can be included without changing the required number of qubits. We provide an example in which χ^2 is nonconvex and show that QFitter can find the global minimum more accurately than its classical alternatives.

DOI: [10.1103/PhysRevD.107.015023](https://doi.org/10.1103/PhysRevD.107.015023)

I. INTRODUCTION

Determining model parameters by adjusting a theoretical hypothesis until it best fits experimental data is commonly referred to as *fitting*. Fitting algorithms use optimization methods that aim to find the model's parameters that result in the smallest value of a characteristic function, like the negative log-likelihood function, χ^2 , which incorporates all experimental observables of interest simultaneously. As such procedure allows an interpretation of experimental data in terms of a given theoretical model, fitting algorithms are at the core of phenomenological studies of new physics.

Effective field theories (EFTs) provide a model-independent framework to tension experimental data with the imprints of new physics scenarios, assuming that any new particles are significantly heavier than the Standard Model particles. The free parameters of an EFT are the coefficients of the local operators that appear in its Lagrangian. They are known as the *Wilson coefficients*. The EFT formulation ensures that the Wilson coefficients

can capture any effects of the heavy particles at low energies compared to their masses.

Because of this, EFT fits have become an essential tool in interpreting experimental data in the current context, in which no resonant new physics has been observed at the LHC. Subsets of the coefficients of the EFT for the nonlinear realization of the electroweak symmetry have been fitted to the relevant experimental data in Refs. [1,2]. Similarly, fits for the linear realization have been performed in Refs. [3–10], for example.

If nonlinear contributions of the Wilson coefficients are considered in the fit, then the χ^2 function to be minimized may develop local minima which differ from the global one. Furthermore, local minima can also occur when nuisance parameters are included, encoding the dependence of the χ^2 function on systematic uncertainties. This may prevent classical optimization algorithms from reaching the values of the Wilson coefficients that fit the observed experimental values optimally. Indeed, such algorithms usually obtain the solution by incrementally improving a given point in parameter space, which can lead to them becoming trapped in a local minimum.

Quantum annealing provides an optimization framework with the potential to perform better than classical algorithms in minimizing nonconvex functions [11–15]. The availability of physical quantum annealing devices with thousands of qubits has made it possible to apply this approach to real-world problems in recent years [16–28].

^{*}juan.c.criado@durham.ac.uk

[†]michael.spannowsky@durham.ac.uk

[‡]roman.kogler@desy.de

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

In this work, we construct a formulation for implementing the problem of fitting the Wilson coefficients of an EFT on these devices. An arbitrary number of observables can be included in the fit without changing the required number of qubits. Furthermore, general polynomial dependence of the theoretical predictions on the Wilson coefficients can be implemented in this formulation, with nonlinear terms encoded by auxiliary qubits. In practice, we find that current quantum annealing devices can be used to perform fits with at least eight Wilson coefficients, including a quadratic dependence of observables in these coefficients. Using practical examples, we show that the quantum approach generates the best-fit estimator for the coefficients accurately and more consistently than classical methods in a nonconvex problem.

The rest of this paper is organized as follows. In Sec. II, we present the QFitter method as one solution to a general fitting problem, and we describe a new formalism needed to apply quantum annealing to the problem of fitting EFT coefficients. In Sec. III, we provide the results of applying QFitter to three different problems, including a fit in which the χ^2 function to be minimized is nonconvex. We compare QFitter to several classical approaches. We give our conclusions in Sec. IV.

II. METHOD

A. Quantum annealing

Quantum annealing is a method for finding the ground state of a given Hamiltonian H_1 . The Hamiltonian of a quantum annealing device takes the form

$$H = A(s)H_0 + B(s)H_1, \quad (1)$$

where s is a free parameter that can be controlled externally; $A(s)$ and $B(s)$ are continuous functions such that $A(0) > 0 = B(0)$ and $A(1) = 0 < B(1)$; and H_0 is a Hamiltonian whose ground state is known in advance.

The solution to the problem is obtained by preparing the system in the ground state of H_0 and changing s continuously from $s = 0$ at an initial time t_i to $s = 1$ at a final time t_f . The function $s(t)$ described by the time evolution of s is known as the *schedule*. The adiabatic theorem ensures that, if the change in s is sufficiently slow, the system is likely to end in the ground state of the target Hamiltonian H_1 .

A concrete realization of this method is transverse-field quantum annealing, which has been implemented in real-world devices. In this realization, the system can be viewed effectively as a collection of qubits (that is, quantum systems with two independent states), with the Hamiltonians H_0 and H_1 given by

$$H_0 = \sum_i \hat{\sigma}_x^i, \quad H_1 = \sum_i h_i \hat{\sigma}_z^i + \sum_{ij} J_{ij} \hat{\sigma}_z^i \otimes \hat{\sigma}_z^j, \quad (2)$$

where h_i and J_{ij} are adjustable parameters, and $\hat{\sigma}_{x,z}^i$ are the x, z Pauli matrices acting on the i th qubit. To perform a computation using transverse-field quantum annealing, one needs to encode its result as the ground state of the Ising Hamiltonian H_1 .

The D-wave devices provide a physical implementation of this setup, in which not all J_{ij} couplings can be set to a nonvanishing value. In the state-of-the-art ADVANTAGE_SYSTEM architecture, there are more than 5000 available qubits, but each one is coupled only to 15 others. To find the ground state of Ising models with a higher degree of connectivity, several qubits are chained together with large coupling to act as a single qubit with more connections. The mapping between the abstract Ising model Hamiltonian to be minimized and the one implemented in the physical device is known as an *embedding*.

Once an embedding has been found, and the schedule $s(t)$ and h_i, J_{ij} parameters are set, the annealer is typically run several times to reduce the effects of external noise. Then, the final state with the least energy obtained from the different runs is selected. The number of runs is referred to in the context as the *number of reads*.

B. Quadratic unconstrained binary optimization formulation

The eigenstates of the quantum Ising Hamiltonian H_1 correspond to the states of its classical analog, whose Hamiltonian is

$$H_{\text{classical}} = \sum_i h_i \sigma_i + \sum_{ij} J_{ij} \sigma_i \sigma_j, \quad (3)$$

with the σ_i being classical variables taking the values $\sigma_i = \pm 1$. Thus, the problem solved by the transverse-field quantum annealers can be viewed equivalently as finding the set of values for the σ_i such that $H_{\text{classical}}$ is minimized:

This problem can be solved both using quantum annealing and classical algorithms, such as simulated annealing. Quantum annealing has been shown to be more consistent in finding the ground state of some nonconvex functions [14]. In Sec. III, we will compare the performance of quantum annealing to several classical methods for EFT fits.

A useful reformulation of the classical Ising Hamiltonian minimization problem is obtained by making use of the binary variables $\tau_i = (\sigma_i + 1)/2$, whose possible values are 0 or 1. In terms of them, the problem can be expressed as the minimization of a homogeneous quadratic polynomial:

$$\min_{\tau_i=0,1} \tau_i Q_{ij} \tau_j. \quad (4)$$

This is known as a quadratic unconstrained binary optimization (QUBO) problem. We will refer to the function $\mathcal{L} = \tau_i Q_{ij} \tau_j$ to be minimized in such a problem as the *loss function*.

C. Log-likelihood as a QUBO

We now tackle the task of finding a QUBO formulation for fits of EFT Wilson coefficients to observables. Let $O_i^{(\text{exp})}$ be the experimentally measured values of the observables under consideration, and $O_i^{(\text{th})}(c)$ the corresponding theoretical predictions, as functions of the collection of Wilson coefficients $c = (c_1, \dots, c_M)$. We assume a Gaussian likelihood $L \propto e^{-\chi^2/2}$, with

$$\chi^2 = \sum_{ab} V_a C_{ab}^{-1} V_b, \quad V_a = O_a^{(\text{exp})} - O_a^{(\text{th})}(c), \quad (5)$$

where C^{-1} is the inverse covariance matrix. The maximum-likelihood estimator for the coefficients can be obtained by minimizing χ^2 .

In any EFT, the theoretical predictions $O_a^{(\text{th})}(c)$ are computed as a series in inverse powers of the cutoff scale Λ , which is cut at a fixed power, depending on the target precision of the calculation. Each Wilson coefficient is associated with an inverse power of Λ . The $O_a^{(\text{th})}(c)$ will thus be polynomials in c . For simplicity, we restrict ourselves here to quadratic polynomials

$$O_a^{(\text{th})}(c) = A_a + \sum_i B_{ai} c_i + \sum_{ij} C_{aij} c_i c_j, \quad (6)$$

although the method that we present can be straightforwardly extended to higher degrees. Nevertheless, quadratic polynomials are sufficient for most current EFT applications. In particular, dimension-eight and dimension-six squared contributions in the Standard Model EFT can be dealt with in this setting.

We thus have a χ^2 function to be minimized that is a quartic polynomial in real variables c_i . To turn this into a QUBO, we need to reformulate it as a quadratic polynomial in binary variables. Binarization can be achieved by means of a binary encoding [15,29]:

$$c_i = L_i + \frac{U_i - L_i}{1 - 2^{-n-1}} \sum_{n=1}^N \frac{c_i^{(n)}}{2^{n+1}}, \quad (7)$$

with L_i , U_i , and N being fixed parameters, and the $c_i^{(n)}$ taking binary values 0 or 1. The coefficients c_i can take 2^N different values, uniformly distributed in the $[L_i, U_i]$ interval. Substituting this in Eq. (6) turns $O_a^{(\text{th})}(c)$ into a quadratic function of the binary variables, which means that χ^2 is a quartic polynomial in them.

The reduction from a fourth-degree polynomial to a second-degree one can be made by means of auxiliary binary variables $c_{ij}^{(m,n)}$, representing the products $c_i^{(m)} c_j^{(n)}$. In this way, the $c_i c_j$ factor that appears in the last term of Eq. (6), can be written as a linear function in the binary variables

$$c_i c_j = L_i L_j + \left[L_j \frac{U_i - L_i}{1 - 2^{-n-1}} \sum_{n=1}^N \frac{c_i^{(n)}}{2^{n+1}} + (i \leftrightarrow j) \right] + \frac{(U_i - L_i)(U_j - L_j)}{(1 - 2^{-n-1})^2} \sum_{m,n=1}^N \frac{c_{ij}^{(m,n)}}{2^{m+n+2}}. \quad (8)$$

Replacing Eq. (7) in the linear term of Eq. (6), and Eq. (8) in the quadratic one, $O_a^{(\text{th})}(c)$ and χ^2 become linear and quadratic in the binary variables, respectively.

In order for this procedure to work, the constraints $c_{ij}^{(m,n)} = c_i^{(m)} c_j^{(n)}$ need to be enforced somehow. This can be done by means of a *constraint Hamiltonian*: a quadratic function $P(x, y, z)$ of binary variables x, y, z that achieves its minimum $P = 0$ if and only if $xy = z$. In particular, we use the function

$$P(x, y, z) = xy - 2z(x + y) + 3z. \quad (9)$$

Now, to construct the loss function, we add together χ^2 , viewed as a quadratic function of the binary variables, as described above, and the constraint Hamiltonians for all the relations between them:

$$\mathcal{L} = \chi^2 + \lambda \sum_{ijmn} P(c_i^{(m)}, c_j^{(n)}, c_{ij}^{(m,n)}). \quad (10)$$

When the coefficient λ is large enough, the minimum of \mathcal{L} will correspond to the minimum of χ^2 over the set of values of the binary variables that satisfy the constraints. This concludes the reformulation of the problem as QUBO: the maximum likelihood estimator for the Wilson coefficients can be obtained by minimizing a function quadratic function of binary variables \mathcal{L} .

The QUBO problem we have obtained can be solved by the usual methods, including quantum and simulated annealing. We will show in Sec. III that, in practice, quantum annealing is the most effective of the two for this purpose, especially when χ^2 is a nonconvex function of the coefficients c_i . The result of any of these methods will be a set of values of the binary variables $c_i^{(n)}$ and $c_{ij}^{(m,n)}$ representing the ground state of \mathcal{L} . The fitted values of the c_i can be recovered using Eqs. (7) and (8).

The number n_{bin} of binary variables employed in this formulation is highly relevant for the practical applications using quantum annealing devices since, together with the connectivity of the QUBO, it will control the number of qubits in the final embedding. n_{bin} is controlled by the number M of Wilson coefficients involved in the fit and the number N of binary variables used in the encoding for each of them. In general, one has

$$n_{\text{bin}} \leq NM + \frac{1}{2}(N+1)NM^2. \quad (11)$$

The first term is the number of binary variables required to encode the coefficients. The second one is for their products: there are $N(N+1)/2$ different combinations of the form $c_i c_j$, and at most M^2 different choices for m and n given (i, j) . In many situations, only certain combinations of coefficients appear in the quadratic terms in Eq. (6). In the examples of Sec. III, only the squares of some individual coefficients appear. Then, the number of products of binary variables to encode is $NM(M-1)/2$, and

$$n_{\text{bin}} \leq NM(M+1)/2. \quad (12)$$

Concerning the connections, an entry of the Q matrix of the QUBO is nonvanishing if and only if there is one observable to which the corresponding coefficients (or product of coefficients) contribute simultaneously. Since there are observables to which many coefficients may contribute, the degree of connectivity is typically high.

Finally, notice here that the number of observables does not play any role in determining n_{bin} . That means that the number of observables that can be included in the fit is not limited by the available number of qubits when the QUBO is embedded in physical annealing devices.

D. Zooming

As discussed in Sec. II A, the total number of available qubits in state-of-the-art devices is in the order of several thousand. However, for problems with a high degree of connectivity that require embedding several physical qubits per abstract qubit in the original problem, the effective number of available qubits is much lower, reaching around 100 for fully connected QUBOs. In an EFT fit, the number M of coefficients is fixed, so the only way to reduce the number of qubits, according to Eq. (11) is to reduce the number N of binary variables per coefficient. Thus, the bound in the number of qubits translates into a bound in N , which limits the precision that can be achieved for the best-fit values of the coefficients.

We overcome this limitation using a zooming process, similar to the one presented in Ref. [30]. The idea is to perform the fit in several steps, which we call *epochs*. Each epoch consists of a quantum annealing run for a refined version of the QUBO problem of the previous run, in which the range $[L_i, U_i]$ for each coefficient c_i is updated. Given a parameter $0 < f < 1$, the *zoom factor*, the update rule for the lower and upper limits of the range is

$$L_i \rightarrow c_i - \frac{f}{2}(U_i - L_i), \quad U_i \rightarrow c_i + \frac{f}{2}(U_i - L_i), \quad (13)$$

where c_i denotes the value of the corresponding coefficient obtained from the previous epoch. The effect of this transformation is to reduce the length of the $[L_i, U_i]$ range by a factor f , while centering it around c_i .

The zooming process thus allows us to achieve any desired precision at the price of introducing a classical update step between quantum annealing runs. We stress that its use is necessary for embedding in current quantum annealing devices because of the limited amount of available qubits. In future devices with a larger quantity of qubits, it might be possible to use a larger number N of binary variables per coefficient, thus making it possible to reach the relevant precision in a single quantum annealing run.

III. RESULTS AND DISCUSSION

A. EWPO fit

As a first test of the QFitter method, we perform a fit of the S and T oblique parameters to the electroweak precision observables. This serves as a validation that the minimization of a convex χ^2 function through quantum annealing can reproduce the results of traditional methods, relying on deterministic approaches. We use the set of observables defined Ref. [31], with the values, uncertainties and correlations given in Refs. [32–36]. The S and T parameters can be defined in this context as

$$S = \frac{16\pi v^2}{gg'} c_{\phi WB}, \quad T = \frac{2\pi(g^2 + (g')^2)v^2}{g^2(g')^2} c_{\phi D}, \quad (14)$$

where v is the Higgs vacuum expectation value, g and g' are the gauge coupling constants for the $SU(2)$ and $U(1)$ factors of the electroweak symmetry, and $c_{\phi WB}$ and $c_{\phi D}$ are the Wilson coefficients for the following dimension-six operators in the Standard Model EFT:

$$\mathcal{O}_{\phi WB} = (\phi^\dagger \sigma^a \phi) W_{\mu\nu}^a B^{\mu\nu}, \quad \mathcal{O}_{\phi D} = |\phi^\dagger D_\mu \phi|^2. \quad (15)$$

Here, ϕ is the Higgs doublet, $W_{\mu\nu}^a$ and $B_{\mu\nu}$ are the $SU(2)$ and $U(1)$ field-strength tensors, D_μ is the corresponding covariant derivative and σ^a are the Pauli matrices.

We choose the following hyperparameters for the fit: $N = 5$ binary variables per coefficient, ranges for them given by $c_{\phi WB} \in [-0.005, 0.005]$ TeV^{-2} and $c_{\phi D} \in [-0.03, 0.03]$ TeV^{-2} , no zooming, 100 reads and a linear anneal schedule $s(t)$, with annealing time $t_f - t_i = 50$ μs . A visual representation of the Q matrix of the corresponding QUBO problem is shown in Fig. 1. We implement this in the D-wave ADVANTAGE_SYSTEM4.1 quantum annealer. The total computing time spent in the quantum annealer for the fit is thus 5 ms. Since there is no zooming procedure involved, the fit is performed purely quantum mechanically in this case, with no classical updating steps. As the result, we obtain the central values $S = 0.05$ and $T = 0.09$, close to the values obtained in similar fits in the literature [37].

As the next step, we generate the profile of the minimal $\Delta\chi^2 = \chi^2 - \chi_{\text{min}}^2$ as a function of the parameter S . To do so, we perform the fit while keeping the value of S fixed. The ranges of for the coefficients are taken

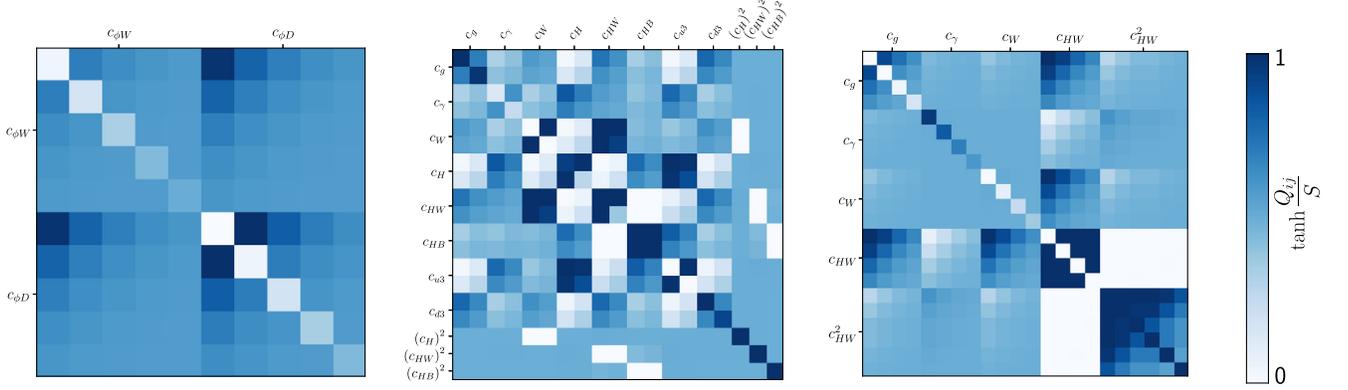


FIG. 1. Qubo matrices Q for the first zooming epoch of the EWPO (left), Higgs (center) and nonconvex (right) fits. The color code shows the value of $\tanh(Q_{ij}/S)$, where $S = 10, 10^2, 10^4$ for the EWPO, Higgs and nonconvex fits, respectively. The lowest value of the matrix is shown in white, while the highest is represented in dark blue.

in this case to be $c_{\phi WB} \in [-0.03, 0.03] \text{ TeV}^{-2}$ and $c_{\phi D} \in [-0.1, 0.04] \text{ TeV}^{-2}$, while the rest of hyperparameters are kept the same as in the fit for the central values. We repeat this for 20 different values of S , equally spaced in a range around its central value. We use the same procedure to produce the profile for $\Delta\chi^2$ as a function of T . The results are shown in Fig. 2. From these profiles, we can derive the following 1σ intervals for the parameters:

$$S = 0.05 \pm 0.09, \quad T = 0.09 \pm 0.08, \quad (16)$$

which are again close to those obtained in Ref. [37]. The total quantum annealing time employed in generating the profiles is 0.2 s.

B. Higgs fit

We now use our method to perform a larger fit, including eight Wilson coefficients and quadratic terms in the theoretical predictions for observables. While this fit can be performed efficiently with classical methods, it serves as a test of the quantum annealing setup to fit more than two parameters simultaneously. The limitations of the current state-of-the-art quantum annealing devices lead to a limited number of parameters that can be determined with QFitter at present. We find that the fit works reliably for up to eight Wilson coefficients. We employ the Higgs fit setup also to verify that the method for reducing quartic polynomials in binary variables to quadratic ones works correctly, and can be used in future fits using quantum annealing. The Lagrangian we consider is a subset of the dimension-six Standard Model EFT Lagrangian, given by

$$\begin{aligned} \mathcal{L} = & \frac{c_{u3}y_t}{v^2} (\phi^\dagger \phi) (\bar{q}_L \tilde{\phi} u_R) + \frac{c_{d3}y_b}{v^2} (\phi^\dagger \phi) (\bar{q}_L \phi d_R) \\ & + \frac{ic_W g}{2m_W^2} (\phi^\dagger \sigma^a D^\mu \phi) D^\nu W_{\mu\nu}^a + \frac{c_H}{4v^2} (\partial_\mu (\phi^\dagger \phi))^2 \\ & + \frac{c_\gamma (g')^2}{2m_W^2} (\phi^\dagger \phi) B_{\mu\nu} B^{\mu\nu} + \frac{c_g g_S^2}{2m_W^2} (\phi^\dagger \phi) G_{\mu\nu}^a G^{a\mu\nu} \\ & + \frac{ic_{HW} g}{4m_W^2} (\phi^\dagger \sigma^a D^\mu \phi) D^\nu W_{\mu\nu}^a \\ & + \frac{ic_{HB} g'}{4m_W^2} (\phi^\dagger D^\mu \phi) D^\nu B_{\mu\nu} + \text{H.c.}, \end{aligned} \quad (17)$$

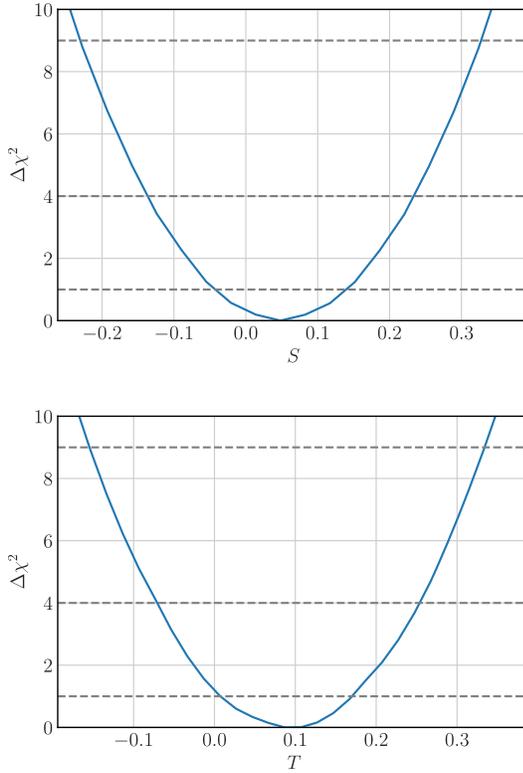


FIG. 2. $\Delta\chi^2 = \chi^2 - \chi^2_{\min}$ profiles for the EWPO fit. For each point in the graphs, χ^2 is minimized while the parameter in the x axis is kept fixed.

where y_t and y_b are the top and bottom Yukawa couplings, m_W is the mass of the W boson and g_s is the strong coupling constant.

We use the set of observables described in Ref. [6]. They correspond to projections of Higgs production and decay processes at the high-luminosity LHC. We consider only inclusive measurements in gluon fusion, vector boson fusion, and the production associated with an electroweak gauge boson, two top quarks or a jet. The decays are into the following final states: $\gamma\gamma$, W^+W^- , ZZ , $\mu^+\mu^-$, $\tau^+\tau^-$ and $b\bar{b}$.

We compute the theoretical predictions for the observables, up to quadratic order in the coefficients. We find that c_H , c_W , c_{HB} and c_{HW} have quadratic contributions to some observables. We only consider contributions of the form c_i^2 for observables where a linear parametrization results in a poor description of the parameter dependence.¹ The only coefficients that have non-negligible contributions at the quadratic level are c_W , c_{HB} and c_{HW} . We assume that the measured values of the coefficients are the same as their (dimension-four) Standard Model values. We use the QUBO formulation of the fit to generate the individual $\Delta\chi^2$ profiles for the coefficients.

There are several choices of hyperparameters that give rise to similar results. However, we find a greater consistency among different trial runs with the following set of values: $N = 2$ binary variables per coefficient, 20 zooming epochs, a zoom factor of $f = 80\%$, 200 reads in each epoch and $t_f - t_i = 100 \mu\text{s}$ of annealing time for each read, with a linear anneal schedule $s(t) \propto t - t_i$. The ranges we use for the coefficients are centered around the origin ($L_i = -U_i$), with the following upper limits:

$$\begin{aligned} U_{u3} &= 2U_{d3} = 10^4 U_g = 10^5 U_\gamma \\ &= U_H = 10U_W = 5U_{HB} = 5U_{HW} = 0.1, \end{aligned} \quad (18)$$

The matrix Q for the corresponding QUBO problem is shown in Fig. 1. Again, we use the D-wave ADVANTAGE_SYSTEM4.1 quantum annealer to perform the fits. The total quantum annealing time per fit is 0.4 s.

We show our results in Fig. 3. For the coefficients c_g , c_γ , c_H , c_{u3} and c_{d3} , we obtain a $\Delta\chi^2$ profile with quadratic shape, which is symmetric under $c_i \rightarrow -c_i$, as only have linear contributions to the observables. For c_W , c_{HB} and c_{HW} , we observe an asymmetry under $c_i \rightarrow -c_i$, generated by the corresponding quadratic terms. Again, we observe good agreement with the results obtained with classical methods [6]. To our knowledge, this is the first time a quantum annealing device has been used to fit eight

¹We parametrize the c_i dependence of each observable by a fit to points generated on a grid and demand $\chi^2 < 0.03$ for a linear parametrization. If $\chi^2 > 0.03$, then we use a quadratic parametrization in c_i .

parameters simultaneously, with some parameters featuring a quadratic dependence in the observables.

C. Nonconvex Higgs fit

Finally, we consider a modified scenario in which the χ^2 function has a local minimum close to $c_i = 0$ and a (much lower) global one away from it. We thus have to minimize nonconvex loss functions, a problem in which the quantum approach might provide an advantage over its classical counterparts [14]. We construct an example in which classical algorithms become trapped in a local minimum and test if the QFitter method can find the global minimum with the same starting values.

To generate this setup, we consider a scenario in which several observables have been measured to be away from their Standard Model values. The observable that plays the most important role in generating the nonconvexity is $\mu_{pp \rightarrow WH \rightarrow WZ\gamma}$, which we set to -0.3 , with total uncertainty to be $\Delta\mu_{pp \rightarrow WH \rightarrow WZ\gamma} = 0.01$. For the rest of the introduced deviations from the Standard Model values, we keep the projected experimental uncertainties while setting the measured values as follows:

- (i) For gluon fusion, vector-boson function, ttH and jH processes with WW or ZZ in the final state, we take $\mu_i = 0.8$.
- (ii) For any process with $Z\gamma$ is in the final state, we set $\mu_i = -6$.
- (iii) For all the processes with Higgs production in association with a vector boson, we choose $\mu_i = 0.2$.

We keep the remaining observables in the fit at their Standard Model values. Concerning the coefficients, we only allow c_{HW} , c_W , c_g and c_γ to vary, while fixing the rest of the coefficients to zero. This simplified scenario leads to a nonconvex negative log-likelihood function. More complex setups are conceivable, but these would need further modifications of the projected observables. We stress that we make these arbitrary choices only to test the performance of QFitter for a nonconvex χ^2 function. The example we have constructed in this way is thus only a toy example which nevertheless contains the crucial property of more realistic examples: the existence of both local and global minima in a multidimensional parameter space. We will use it to show the reliability of QFitter in finding the global minimum in this situation, which we take as a strong indication of its potential to solve more complex optimization problems when larger annealing devices with more qubits become available.

For the fits in this section, we use the QUBO formulation with $N = 4$ binary variables per coefficient. The hyperparameters are chosen the same as in Sec. III B: 20 zooming epochs, a zooming factor $f = 80\%$, 200 reads per epoch, and $t_f - t_i = 100 \mu\text{s}$ annealing time, with a linear schedule $s(t)$. The ranges of values for the coefficients that we choose are $c_g, c_\gamma \in [-2, 2] \times 10^{-5}$, $c_{HW}, c_H \in [-0.1, 0]$.

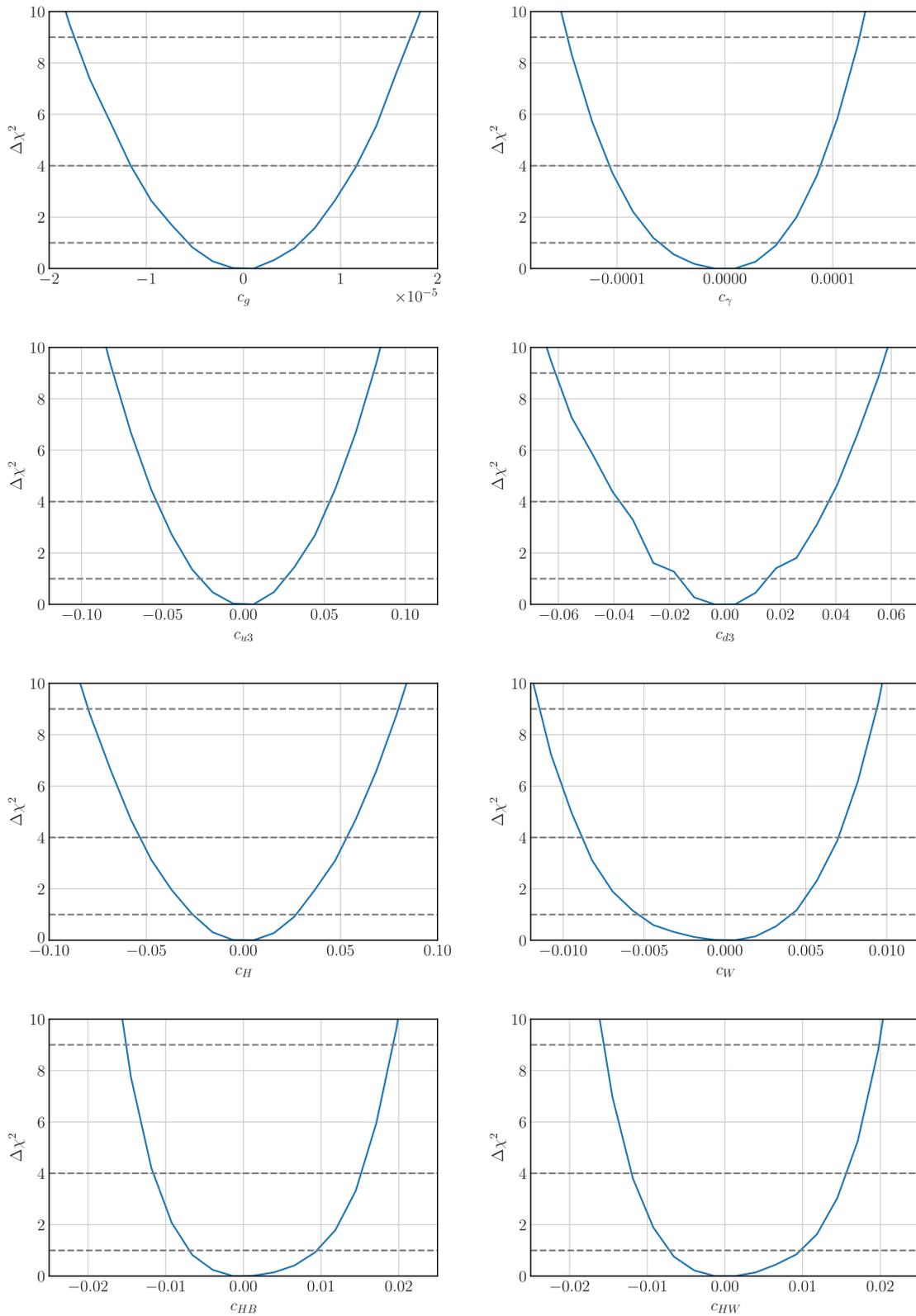


FIG. 3. Minimal $\Delta\chi^2 = \chi^2 - \chi^2_{\min}$ as a function of the each of the Wilson coefficients in the Higgs fit.

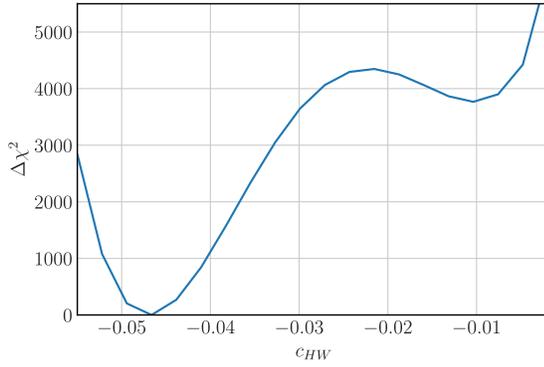


FIG. 4. $\Delta\chi^2 = \chi^2 - \chi_{\min}^2$ as a function of c_{HW} , displaying a local minimum close to $c_{HW} = 0$ and a global one near $c_{HW} = -0.05$.

The Q for the QUBO problem corresponding to the first zooming epoch is shown in Fig. 1.

We embed this QUBO in the D-wave ADVANTAGE_SYSTEM4.1 quantum annealer, fixing c_{HW} at 20 equally spaced values in its range to generate the $\Delta\chi^2$ function when profiling over the other parameters. The results are given in Fig. 4. This suggests that χ^2 indeed has two local minima, one at $c_{HW} \simeq -0.01$ and a lower one at $c_{HW} \simeq -0.05$. We check that the gradients are small at these two points and that the Hessian is definite positive. As a quantitative measure of the smallness of the gradients, we find that the following equation is satisfied:

$$\sum_i \frac{c_i}{10} \frac{\partial_c \chi^2}{\chi^2} < 10\%. \quad (19)$$

We conclude that we have found a good approximation to two of the local minima of χ^2 . As shown in the Appendix, the form of the χ^2 function that we consider here implies that we can have at most two local minima, so we have identified all of them. With $\chi^2 \geq 0$, this implies that the minimum at $c_{HW} \simeq -0.05$ is the global one.

We now compare different methods for finding the maximum-likelihood estimations of the four coefficients. This corresponds to a search for the global minimum of χ^2 . The classical methods that we consider are

- (i) The MIGRAD algorithm provided by the MINUIT code [38]. This type of variable-metric algorithm works by incrementally improving an approximation of the error matrix and the best-fit point.
- (ii) Simulated annealing. This method also proceeds by iteratively improving a point in parameter space. In each step, the algorithm moves to a random nearby point if the new point has a lower χ^2 . If the new point has a higher χ^2 , then it moves to it with probability $e^{-\Delta\chi^2/T}$, where $\Delta\chi^2$ is the difference in χ^2 between the new and the old points, and T is a parameter that decreases along the run, known as the temperature. The algorithm explores the vicinity of local minima

while changing T depending on the number of failures to find a better minimum in the previous steps. This may allow the algorithm to escape local minima in some cases. We chose the initial value of T to be 10^6 with a minimum value of 10^{-6} , an adaptive speed of 1 and a step size of 0.01. The maximum number of steps is chosen to be 10^5 .

- (iii) Finally, we also consider the same algorithm as for the quantum annealing approach, with the QUBO formulation and a zooming process, replacing the quantum annealing runs with simulated annealing. In this case, the parameter points to be updated by the algorithm are the sets of 0 or 1 values of the binary variables. The random nearby point is obtained by flipping a random variable. The number of steps is usually measured here in terms of *sweeps*, which correspond to as many updating steps as binary variables are present in the problem.

We refer to the methods not using the QUBO formulation as *standard-formulation* methods. We show a comparison of the best-fit values of the coefficients and χ^2 obtained from all the methods in Table I. The classical methods are initialized to a point with $c_H = c_g = c_\gamma = 0$.

Inspecting the χ^2 values given in the Table I, one can see that classical algorithms that start at the point $c_i = 0$ typically get trapped in the local minimum nearby. The MIGRAD algorithm always finds the local minimum at $c_{HW} \simeq -0.009$, when the starting value of c_{HW} is 0. Even the standard simulated annealing algorithm, which should perform better in rejecting local minima, gets trapped in this minimum. We have observed this to be the case despite setting the initial value of c_{HW} to -0.05 . We observe that the other three free parameters have a large impact on the behavior of the fit and impair the ability of the algorithm to find the global minimum. This problem is not present in the quantum annealing approach. The minimum with $\chi^2 = 135$ obtained by MINUIT for initial point $c_{HW} = -0.05$ corresponds to the one from the quantum annealer. The small differences in the minimum value of χ^2 and the best-fit parameters are because we have neglected some quadratic contributions in the formulation of the QUBO problem, which has no consequences on the general shape of the χ^2 function, with a global minimum around $c_{HW} = -0.05$ and a local one close to $c_{HW} = 0$.

The QUBO-formulation simulated annealing results depend strongly on the schedule for the temperature. For large starting temperatures, there is no dependence on the starting point. The results also vary considerably for fixed annealing parameters from run to run. We find the most consistent results with a schedule that exponentially increases the inverse-temperature parameter β from 1×10^{-5} to 10 in 1×10^6 steps, performing one sweep per step. The results fall into three classes: (A) those on the “wrong” side of the barrier, with $\chi^2 \simeq 4000$; (B) those on

TABLE I. Best fit values of the Wilson coefficients from different fitting methods, for the nonconvex χ^2 Higgs fit. The QUBO-formulation simulated annealing generates radically different results for different runs with the same parameters, which fall into three classes: (A) $\chi^2 \simeq 4000$, (B) $\chi^2 < 1000$ and (C) $\chi^2 > 10^7$, occurring 13%, 82% and 5% of the time, respectively. We present here an example of class-A and an example class-B results. The quantum annealing results shown correspond to the smallest value of χ^2 across 10 runs. We find an average minimum χ^2 value of 70.8 with a standard deviation of 2.4.

Method		Fit time (s)	c_{HW}	c_H	c_g	c_γ	χ^2
Standard	MINUIT (initial $c_{HW} = 0$)	2.0	-0.009	0.100	1.4×10^{-5}	3.2×10^{-6}	4110
	MINUIT (initial $c_{HW} = -0.05$)	2.4	-0.050	0.039	-9.7×10^{-6}	-1.0×10^{-4}	135
	Simulated annealing (initial $c_{HW} = 0$)	642	-0.009	0.100	1.4×10^{-5}	3.7×10^{-6}	4110
	Simulated annealing (initial $c_{HW} = -0.05$)	644	-0.009	0.100	1.4×10^{-5}	3.7×10^{-6}	4110
QUBO	Simulated annealing (Class A)	6.4	-0.012	-0.054	-3.0×10^{-5}	3.9×10^{-5}	3910
	Simulated annealing (Class B)	6.4	-0.045	-0.175	-3.7×10^{-5}	1.8×10^{-4}	228
	Quantum annealing	0.2	-0.047	-0.050	1.9×10^{-5}	7.5×10^{-7}	68

the correct side, with $\chi^2 < 1000$; and (C) those for which the zooming gets stuck at $\chi^2 > 10^7$. We perform 40 runs and find that 13% of the runs belong to class A, 82% to class B and 5% to class C. The results in class (B) also present a considerable variation. We show an arbitrary example in Table I.

In order to test the performance of QFitter in finding the global minimum, we perform 10 independent runs, keeping c_{HW} free. The fitted value of c_{HW} is consistently $\simeq -0.05$ in all the runs, implying that QFitter does not get trapped close to $c_{HW} \simeq -0.01$. We find some variation in the minimum χ^2 values across the runs: five fall in the interval [68, 70], four in the interval [70, 72], and one lies at a value of 78. These variations originate from the finite time spent by the algorithm on the quantum annealing device. The exact ground state is only guaranteed to be reached for an infinite amount of time. However, the algorithm always results in a point in parameter space close to the global minimum. In practice, one could run the MIGRAD algorithm after the QFitter method to obtain the exact minimum.

In Table I, we also provide the total time spent performing the optimization for each method. Again, we run MINUIT, the standard-formulation simulated annealing and the QUBO simulated annealing in an Apple M1 processor. Since quantum annealing is performed on a dedicated device, the numbers cannot be compared directly. However, we note that quantum annealing requires orders of magnitude less time to perform this task.

IV. CONCLUSIONS

We have presented QFitter, a quantum annealing-based method for fitting EFT coefficients to experimental measurements. The χ^2 is encoded as a QUBO problem which can be directly embedded in the currently available quantum annealers. The required number of qubits depends on the number of coefficients to be fitted, the nonlinear terms in their contributions to observables (which require auxiliary qubits), and the precision to which they are to

be determined. The number of observables included does not affect this.

Since physical annealers only provide a limited amount of qubits, the practical implementation of QFitter can only be done for a limited number of coefficients, precision and nonlinearities. We have used a zooming algorithm, in which the precision is increased iteratively through several annealing runs, to overcome this limitation partially. With this setup, we have found that fitting problems involving at least eight coefficients and their quadratic dependencies can be embedded in current quantum annealing devices.

Finally, we have tested the performance of QFitter with three examples. The first two, the EWPO and the Higgs fit, involve a convex χ^2 function. The quantum approach gives comparable results to the classical ones here. We have then modified the χ^2 for the Higgs fit to make it nonconvex. By comparing with several classical algorithms, we have found that the quantum one is the one that ends in the global minimum most consistently with a considerable gain in processing time.

APPENDIX: NUMBER OF LOCAL MINIMA IN THE NONCONVEX HIGGS FIT

We prove here that the χ^2 function defined in Sec. III C has two local minima at the most. Neglecting constant terms, χ^2 can be written as

$$\chi^2 = R_{ij}x_i x_j + S_{ij}y_i y_j + T_{ij}x_i y_j, \quad (\text{A1})$$

where the $R_{ij} = R_{ji}$, $S_{ij} = S_{ji}$ and T_{ij} are constant parameters, and $x = (c_W, c_g, c_\gamma)$, $y = (c_{HW}, c_{HW}^2)$. It can be checked that the matrix R is invertible. At a local minimum we must have $\nabla \chi^2 = 0$. In particular, $\partial_{x_i} \chi^2 = 0$ must be satisfied. That is,

$$2R_{ij}x_j = -T_{ij}y_j. \quad (\text{A2})$$

This is a system of linear equations in x_j with a unique solution $\bar{x}_i(y) = -\frac{1}{2}R_{ik}^{-1}T_{kj}y_j$. Since Eq. (A2) is a necessary

condition for a point (x, y) to be a local minimum, all the local minima of χ^2 must lie on the curve described by $(\bar{x}(y), y)$. We can thus look for the local minima of

$$\chi^2|_{x_i=\bar{x}_i(y)} = R_{ij}\bar{x}_i(y)\bar{x}_j(y) + S_{ij}y_i y_j + T_{ij}\bar{x}_i(y)y_j. \quad (\text{A3})$$

The fact that $\bar{x}_i(y)$ is linear in y implies that it is quadratic in c_{HW} . It follows that $\chi^2|_{x_i=\bar{x}_i(y)}$ is a quartic polynomial in c_{HW} . This concludes the proof, since a quartic polynomial can only have up to two local minima.

-
- [1] G. Buchalla, O. Cata, A. Celis, and C. Krause, Fitting Higgs data with nonlinear effective theory, *Eur. Phys. J. C* **76**, 233 (2016).
- [2] I. Brivio, J. Gonzalez-Fraile, M. C. Gonzalez-Garcia, and L. Merlo, The complete HEFT Lagrangian after the LHC run I, *Eur. Phys. J. C* **76**, 416 (2016).
- [3] C. Englert, R. Kogler, H. Schulz, and M. Spannowsky, Higgs coupling measurements at the LHC, *Eur. Phys. J. C* **76**, 393 (2016).
- [4] A. Buckley, C. Englert, J. Ferrando, D. J. Miller, L. Moore, M. Russell, and C. D. White, Constraining top quark effective theory in the LHC run II era, *J. High Energy Phys.* **04** (2016) 015.
- [5] M. Klute, R. Lafaye, T. Plehn, M. Rauch, and D. Zerwas, Measuring Higgs Couplings from LHC Data, *Phys. Rev. Lett.* **109**, 101801 (2012).
- [6] C. Englert, R. Kogler, H. Schulz, and M. Spannowsky, Higgs characterisation in the presence of theoretical uncertainties and invisible decays, *Eur. Phys. J. C* **77**, 789 (2017).
- [7] J. Ellis, C. W. Murphy, V. Sanz, and T. You, Updated global SMEFT fit to Higgs, Diboson and electroweak data, *J. High Energy Phys.* **06** (2018) 146.
- [8] J. J. Ethier, G. Magni, F. Maltoni, L. Mantani, E. R. Nocera, J. Rojo, E. Slade, E. Vryonidou, and C. Zhang (SMEFT Collaboration), Combined SMEFT interpretation of Higgs, diboson, and top quark data from the LHC, *J. High Energy Phys.* **11** (2021) 089.
- [9] J. de Blas, Y. Du, C. Grojean, J. Gu, V. Miralles, M. E. Peskin, J. Tian, M. Vos, and E. Vryonidou, Global SMEFT fits at future colliders, [arXiv:2206.08326](https://arxiv.org/abs/2206.08326).
- [10] Anisha, S. Das Bakshi, S. Banerjee, A. Biekötter, J. Chakraborty, S. Kumar Patra, and M. Spannowsky, Effective limits on single scalar extensions in the light of recent LHC data, [arXiv:2111.05876](https://arxiv.org/abs/2111.05876).
- [11] E. Farhi, J. Goldstone, and S. Gutmann, Quantum adiabatic evolution algorithms versus simulated annealing, [arXiv:quant-ph/0201031](https://arxiv.org/abs/quant-ph/0201031).
- [12] S. Abel, N. Chancellor, and M. Spannowsky, Quantum computing for quantum tunneling, *Phys. Rev. D* **103**, 016008 (2021).
- [13] S. Abel and M. Spannowsky, Observing the fate of the false vacuum with a quantum laboratory, *PRX Quantum* **2**, 010349 (2021).
- [14] S. Abel, A. Blance, and M. Spannowsky, Quantum optimisation of complex systems with a quantum annealer, *Phys. Rev. A* **106**, 042607 (2022).
- [15] S. Abel, J. C. Criado, and M. Spannowsky, Completely quantum neural networks, *Phys. Rev. A* **106**, 022601 (2022).
- [16] A. B. Finilla, M. A. Gomez, C. Sebenik, and J. D. Doll, Quantum annealing: A new method for minimizing multidimensional functions, *Chem. Phys. Lett.* **219**, 343 (1994).
- [17] T. Kadowaki and H. Nishimori, Quantum annealing in the transverse Ising model, *Phys. Rev. E* **58**, 5355 (1998).
- [18] J. Brooke, D. Bitko, T. F. Rosenbaum, and G. Aeppli, Quantum annealing of a disordered magnet, *Science* **284**, 779 (1999).
- [19] N. G. Dickson *et al.*, Thermally assisted quantum annealing of a 16-Qubit problem, *Nat. Commun.* **4**, 1903 (2013).
- [20] T. Lanting *et al.*, Entanglement in a Quantum Annealing Processor, *Phys. Rev. X* **4**, 021041 (2014).
- [21] T. Albash, W. Vinci, A. Mishra, P. A. Warburton, and D. A. Lidar, Consistency tests of classical and quantum models for a quantum annealer, *Phys. Rev. A* **91**, 042314 (2015).
- [22] T. Albash and D. A. Lidar, Adiabatic quantum computing, *Rev. Mod. Phys.* **90**, 015002 (2018).
- [23] S. Boixo, V. N. Smelyanskiy, A. Shabani, S. V. Isakov, M. Dykman, V. S. Denchev, M. H. Amin, A. Y. Smirnov, M. Mohseni, and H. Neven, Computational multiqubit tunneling in programmable quantum annealers, *Nat. Commun.* **7**, 10327 (2016).
- [24] N. Chancellor, S. Szoke, W. Vinci, G. Aeppli, and P. A. Warburton, Maximum-entropy inference with a programmable annealer, *Sci. Rep.* **6**, 22318 (2016).
- [25] M. Benedetti, J. Realpe-Gómez, R. Biswas, and A. Perdomo-Ortiz, Estimation of effective temperatures in quantum annealers for sampling applications: A case study with possible applications in deep learning, *Phys. Rev. A* **94**, 022308 (2016).
- [26] S. Muthukrishnan, T. Albash, and D. A. Lidar, Tunneling and Speedup in Quantum Optimization for Permutation-Symmetric Problems, *Phys. Rev. X* **6**, 031010 (2016).
- [27] A. Cervera Lierta, Exact Ising model simulation on a quantum computer, *Quantum* **2**, 114 (2018).
- [28] T. Lanting, The d-wave 2000q processor, in *Proceedings of the AQC 2017* (2017).
- [29] J. C. Criado and M. Spannowsky, Qade: Solving differential equations on quantum annealers, *Quantum Sci. Technol.* **8**, 015021 (2023).

- [30] A. Zlokapa, A. Mott, J. Job, J.-R. Vlimant, D. Lidar, and M. Spiropulu, Quantum adiabatic machine learning by zooming into a region of the energy surface, *Phys. Rev. A* **102**, 062405 (2020).
- [31] J. Aebischer, J. Kumar, P. Stangl, and D.M. Straub, A global likelihood for precision constraints and flavour anomalies, *Eur. Phys. J. C* **79**, 509 (2019).
- [32] S. Schael *et al.* (ALEPH, DELPHI, L3, OPAL, SLD Collaborations, LEP Electroweak Working Group, SLD Electroweak Group, and SLD Heavy Flavour Group), Precision electroweak measurements on the Z resonance, *Phys. Rep.* **427**, 257 (2006).
- [33] S. Schael *et al.* (ALEPH, DELPHI, L3, OPAL and LEP Electroweak Collaborations), Electroweak measurements in electron-positron collisions at W-Boson-pair energies at LEP, *Phys. Rep.* **532**, 119 (2013).
- [34] T. A. Aaltonen *et al.* (CDF and D0 Collaborations), Combination of CDF and D0 W-Boson mass measurements, *Phys. Rev. D* **88**, 052018 (2013).
- [35] C. Patrignani *et al.* (Particle Data Group), Review of particle physics, *Chin. Phys. C* **40**, 100001 (2016).
- [36] M. Aaboud *et al.* (ATLAS Collaboration), Measurement of the W-boson mass in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector, *Eur. Phys. J. C* **78**, 110 (2018); Erratum, *Eur. Phys. J. C* **78**, 898 (2018).
- [37] J. Haller, A. Hoecker, R. Kogler, K. Mönig, T. Peiffer, and J. Stelzer, Update of the global electroweak fit and constraints on two-Higgs-doublet models, *Eur. Phys. J. C* **78**, 675 (2018).
- [38] F. James and M. Roos, MINUIT: A system for function minimization and analysis of the parameter errors and correlations, *Comput. Phys. Commun.* **10**, 343 (1975).