Efficient Prediction of Grain Boundary Energies from Atomistic Simulations via Sequential Design

Martin Kroll, Timo Schmalofski, Holger Dette,* and Rebecca Janisch

1. Introduction

Grain boundaries (GBs) play an important role in materials science but they are still among the least understood crystal defects. Nano-structured materials, that is materials with a high density of GBs, dominate the development of new technologies for energy harvesting and storage,[2] for data processing, but also of GBs, dominate the development of new technologies for energy harvesting and storage,[2] for data processing, but also of GBs, dominate the development of new technologies for energy harvesting and storage,[2] for data processing, but also of GBs, dominate the development of new technologies for energy harvesting and storage,[2] for data processing, but also of GBs, dominate the development of new technologies for energy harvesting and storage,[2] for data processing, but also of GBs, dominate the development of new technologies for energy harvesting and storage,[2] for data processing, but also of GBs, dominate the development of new technologies for energy harvesting and storage,[2] for data processing, but also of GBs, dominate the development of new technologies for energy harvesting and storage.[2] 

With the goal of improving data based materials design, it is shown that by a sequential design of experiment scheme the process of generating and learning from the data can be combined to discover the relevant sections of the parameter space. The application is the energy of grain boundaries as a function of their geometric degrees of freedom, calculated from a simple model, or via atomistic simulations. The challenge is to predict the deep cusps of the energy, which are located at irregular intervals of the geometric parameters. Existing sampling approaches either use large sets of datapoints or a priori knowledge of the cusps' positions. By contrast, the authors’ technique can find unknown cusps automatically with a minimal amount of datapoints. Key point is a Kriging interpolator with Matérn kernel to estimate the energy function. Using the jackknife variance, the next point in the sequential design is a compromise between sampling the region of largest fluctuations and avoiding a clustering of datapoints. In this way, the cusps of the energy can be found within only a few iterations, and refined as desired. This approach will be advantageous for any application with strong, localized fluctuations in the values of the unknown function.

M. Kroll, H. Dette
Fakultät für Mathematik
Ruhr-Universität Bochum
Bochum 44780, Germany
E-mail: holger.dette@rub.de
T. Schmalofski, R. Janisch
Interdisciplinary Centre for Advanced Materials Simulation (ICAMS)
Ruhr-Universität Bochum
Bochum 44780, Germany

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/adts.202100615

© 2022 The Authors. Advanced Theory and Simulations published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

DOI: 10.1002/adts.202100615
boundaries with mixed tilt and twist character. Dimensional parameter spaces, for example, the 2D space of grain boundary energies, and the model still represents a rather tall lattice. [32] This paper will exemplarily focus on such 1D STGB subspace with different fixed rotation axes. Bulatov et al. [3] then used the RSW model to fit the energy in the 1D subspaces, and a simple interpolation approach, and combine it with an alternative interpolation method. The combination of these two recipes leads to an algorithm, which can determine the energy cusps with a minimal number of sequential steps and high accuracy.

In Section 2.1, we first compare two interpolation methods for exploring 1D sections of the energy landscape which can be parametrized by one single parameter, usually an angle. The first one is an expansion of the grain boundary energy in a truncated trigonometric series (introduced in Section Truncated Trigonometric Series Expansion), inspired by the use of spherical harmonics in materials science. [1,10,11] The second interpolation method is the so-called Kriging (also referred to as Gaussian process regression in the presence of noisy observations), which is a well-established method in geostatistical applications [20] and has recently attracted interest in the materials science literature. [20] This method predicts the value of a function at a given point by computing a weighted average of the already determined values of the function in the neighborhood of the point. It is somewhat related to a regression analysis, but Kriging exactly interpolates through the observations already made. Our choice of Kriging interpolator is, among other parameters, defined by a smoothness parameter $\nu$ (see Section Simple Kriging Interpolator for more details). The smaller the value of $\nu$, the rougher is the result-}

energy for fixed misorientation axes. To obtain it, one has to prescribe the GB plane. The vector $\hat{r}$ can be parallel to the misorientation axis, in which case the GB is a pure twist GB, or perpendicular, which results in a pure tilt GB, or it can have a mixed character. If the GB is a tilt grain boundary and the plane represents a mirror plane between the two grains, the GB is called a symmetrical tilt grain boundary (STGB). In this case, for each given rotation axis, the misorientation angle is the only DOF left. The energy as a function of misorientation angle for STGBs shows pronounced minima, so-called cusps. They occur at special misorientations, for which the rotation of the two grains leads to a periodic superlattice of coincidence sites across both grains. Such coincidence-site lattice (CSL) based GBs are characterized by their $\Sigma$ value, where $\Sigma = V_{\text{CSL}} / V_0$, the ratio of the volume of a periodic unit of the CSL lattice to the periodic unit of the crystal lattice. [12] This paper will exemplarily focus on such 1D STGB subspaces with different fixed rotation axes.

The so-called Read–Shockley model of GB energies [26] (see Section The Read–Shockley–Wolf Model) with the empirical extension of Wolf [35] (called Read–Shockley–Wolf (RSW) model in the following) provides a good description of the 1D energy subspaces. Olmsted et al. [21] calculated the energies of several tilt, twist and mixed grain boundaries, from atomistic simulations, starting with 1D subspaces and moving further in the parameter space to expand these sets. Bulatov et al. [3] then used the RSW model to fit the energy in the 1D subspaces, and a simple interpolation was applied between them. In this approach, only several hundred datapoints are required for a satisfying fit. However, the overall strategy requires a priori knowledge of the topography of the energy space, and the model still represents a rather crude approximation of the (up to now rarely explored) higher dimensional parameter spaces, for example, the 2D space of grain boundaries with mixed tilt and twist character.

As an alternative strategy, a high-throughput simulation method was employed by Kim et al. [11] to generate a database of more than 68,000 grain boundary energies in bcc iron, for fixed steps in misorientation and inclination angle. While such a regular sampling approach is quick, since it can easily be parallelized, it misses the special $\Sigma$ grain boundaries. In ref. [13] these structures were added to the list of sampled grain boundaries manually. Restrepo et al. [27] used this data to train and test an artificial neural network (ANN) for the prediction of grain boundary energies. 25% of the data, including low-energy, special grain boundaries, was used for the training of the network, 75% for testing. This study showed that a trained ANN is able to predict unknown grain boundary energies with an error of less then 4%. Still, one needs thousands of grain boundaries for the training and additional knowledge about the position of the cusps.

To arrive at a more effective method for sampling the grain boundary parameter space, we replace in this paper the regular sampling technique by a sequential design of experiments approach, and combine it with an alternative interpolation method. The combination of these two recipes leads to an algorithm, which can determine the energy cusps with a minimal number of sequential steps and high accuracy.

In Figure 1b. In practical applications, $\nu$ can be fixed in advance or estimated from the given data. In our study, we have considered both scenarios with respect to the quality of the fit.

An important advantage of Kriging consists in the fact that it comes along with a natural measure of uncertainty quantification for predictions at potential locations for future observations.
We will use this measure to develop a sequential design strategy, which advises the experimenter to put new observations into areas with large uncertainty. The method is explained in detail in Section Sequential Designs for Kriging and computes a criterion function on a set of potential future sampling locations. It measures the uncertainty concerning the target energy function at any given location. The next sampling point is then selected as a maximizer of the criterion function. Altogether, this leads to a procedure where the next design point depends on both the already chosen sampling locations and the corresponding energies.

In the first part of the paper, we use the RSW model as a quick method to generate grain boundary energies as benchmark data. The mathematical details of the model are given in Section The Read–Shockley–Wolf Model. The numerical results obtained for the RSW data in Section 2.1 suggest a better performance of Kriging interpolation compared to series estimation. The Kriging method is then further refined by means of a sequential sampling technique. The key idea of this sampling technique, which is described in detail in Section Sequential Designs for Kriging, is to make observations at regions of interest like cusps more likely. As a consequence the sequentially sampled data gives a better overall picture of the grain boundary energy landscape. Our results in Section 2.2 demonstrate that substantial improvements can be obtained by sequential sampling.

In Section 2.3 we investigate the performance of the Kriging interpolator and the sequential design for atomistic simulations. Grain boundary energies for a reference data set as well as during the sequential experiments are determined via molecular statics with a semi-spherical grain method and an embedded atom method potential (see Section Atomistic Simulation Method for technical details). Using the atomistic data, details of the Kriging method such as the choice of a smoothness parameter are optimized.

Our study demonstrates that, in contrast to the regular sampling techniques, the sequential sampling technique proposed in this paper is able to sample a subspace with a small number of observations and to identify at the same time unknown cusps.

2. Results

2.1. Kriging versus Series Estimation: A Comparison Using RSW Model Data

We begin with a comparison of the two interpolation procedures using data predicted by the RSW model. Note that (in contrast to the atomistic simulations approach presented below) evaluations in the RSW are not computationally expensive because an explicit representation for the energy function is available. Therefore this model provides an efficient tool for testing the potential applicability of the proposed interpolation procedures.

To be precise, we chose the energy function of the [110] symmetrical tilt grain boundaries (where [110] is the rotation axis) in a cubic metal, because it contains several energy cusps and maxima. It is obtained by specifying the energy function, denoted by \( F_{RSW} \), to contain seven logarithmic segments positioned at the misorientation angles and energy offset values (see Section Trigonometric Series and Kriging Interpolation for details). This function is displayed by the dashed line in Figure 1. We now assume that the function is only observed at a sample of \( N = 9, 17, 33 \), and 65 equally spaced points, say \( (\theta_1, F_{RSW}(\theta_1)), \ldots, (\theta_N, F_{RSW}(\theta_N)) \), which are marked by black triangles in Figure 1 for the case \( N = 17 \). We compare the prediction properties of two interpolation methods, which are described in detail in Section Trigonometric Series and Kriging Interpolation: trigonometric series and Kriging. The predictions of both procedures, which interpolate exactly through the observed data points, are indicated by the blue solid lines where the left panel corresponds to series and the right panel to Kriging interpolation.

As the performance measure for a generic interpolator \( \hat{F}_{RSW} \) we use throughout the paper the maximum absolute error

\[
\max_{\theta \in [0, 180]} |\hat{F}_{RSW}(\theta) - F_{RSW}(\theta)|
\]

(1)

between the true function \( F_{RSW} \) and its interpolator \( \hat{F}_{RSW} \) calculated from the sample

\[
(\theta_1, F_{RSW}(\theta_1)), \ldots, (\theta_N, F_{RSW}(\theta_N))
\]

(2)

Here, \( \theta_{max} = 180^\circ \) is the maximum misorientation angle of the considered grain boundary space, taking into account the symmetries of the cubic crystal system.

Corresponding results for the different sample sizes are shown in Table 1 for the trigonometric series and the Kriging interpolator, using different fixed smoothness parameters \( \nu \), as well as a data adaptive choice, which is denoted by \( \nu^* \) and obtained by maximum likelihood estimation (see Section Simple Kriging Interpolator for details). We observe that in all cases but one, the Kriging interpolator outperforms the trigonometric interpolator, independently of the choice of the parameter \( \nu \), or whether \( \nu^* \) was determined with the data-adaptive maximum likelihood estimator. The exception is \( N = 33 \) and \( \nu = 0.5 \), which results in a maximum error of 0.1424, which is larger than the one for trigonometric interpolation, 0.1245 (note that for all other smoothing parameters Kriging outperforms trigonometric interpolation). We have considered numerous other tilt and twist grain boundaries in the RSW model, and observed a similar superiority of the Kriging approach. The results are displayed in Tables A1–A5 in Appendix A.1. From these numerical experiments we conclude that Kriging generally outperforms the series interpolator, and that for data which can be represented within the RSW model, Kriging interpolation with a fixed smoothing parameter \( \nu = 2.5 \) is a very good choice. In the next step, we introduce a sequential design strategy aiming at a further improvement of the Kriging interpolator.

Table 1. Maximum absolute error (1) for trigonometric series and Kriging interpolation. Different values for the parameter \( \nu \) including a data-adaptive choice \( \nu^* \) by means of maximum likelihood estimation (MLE) were considered for the Kriging interpolator.

| \( N \) | Trigonometric series | \( \nu = 0.5 \) | \( \nu = 1.3 \) | \( \nu = 2.5 \) | \( \nu^* \) via MLE |
|---|---|---|---|---|---|
| 9  | 0.3636 | 0.2570 | 0.1924 | 0.1784 | 0.2569 (\( \nu^* = 0.5062 \)) |
| 17 | 0.2117 | 0.1751 | 0.1357 | 0.1317 | 0.1557 (\( \nu^* = 0.6572 \)) |
| 33 | 0.1245 | 0.1424 | 0.1079 | 0.1040 | 0.1090 (\( \nu^* = 1.3726 \)) |
| 65 | 0.0688 | 0.0361 | 0.0335 | 0.0352 | 0.0322 (\( \nu^* = 1.2561 \)) |
2.2. Sequential Designs in the RSW Model

In a second series of experiments we test a sequential design strategy for the Kriging interpolator. Roughly speaking, this procedure has two steps. A certain number of experiments (initial design) is conducted to calculate a first Kriging interpolator. The remaining experiments are carried out according to a sequential design, reflecting the previous output of the interpolator and updating it, such that more interesting regions of the grain boundary energy space are better explored. The details of this sequential procedure are described in Section Sequential Designs for Kriging.

We assume that we have an overall budget of $N_{\text{total}} = 9, 17, 33, 65$ observations. As initial designs we consider regular equidistant grids on the interval $[0, \theta_{\text{max}}]$, where the number of points of the initial design is $N_{\text{init}} = 9, 17, 33, 65$ (note that by this choice the initial designs are nested, which makes them directly comparable). We then conducted $N_{\text{total}} - N_{\text{init}}$ further simulation experiments, where the experimental conditions are chosen according to the sequential design algorithm described in Section Sequential Designs for Kriging. For the smoothness parameter $\nu$ of the Kriging interpolator, we considered here only the value $\nu = 2.5$ which performed very well in the experiments with fixed equidistant start design (see Table 1).

Figure 2a illustrates the dynamics of the sequential design algorithm. It shows that the sequential design approach selects future sampling points alternately at different parts of the sampling domain $[0, \theta_{\text{max}}]$ with a high proportion of points close to the a priori unknown cusps. A comparison of the prediction error of the Kriging interpolator of a non-sequential and sequential approach is summarized in Figure 2b. This figure shows the maximum error for different combinations of initial design sizes $N_{\text{init}}$ and final design sizes $N_{\text{total}}$ after application of the sequential design algorithm (consequently, $N_{\text{seq}} = N_{\text{total}} - N_{\text{init}}$ points were chosen by the sequential design strategy). For instance, the first line corresponds to a total budget of $N_{\text{total}} = 65$ experiments, with four entries corresponding to the sample size $N_{\text{init}} = 9, 17, 33$, and 65 for the initial design (in the last case the are no sequential experiments). In this case the smallest error (namely 0.008) is obtained for a sequential design with $N_{\text{init}} = 17$ equally spaced design points in the starting design and $N_{\text{seq}} = 48$ sequential observations. We observe that the sequential strategy can significantly outperform the non-sequential approach, for which the maximum error is four times larger.

In Figure 3 we display the non-sequential and sequential design, the resulting Kriging interpolator (blue line) together with the absolute error as a function of the misorientation angle. We observe that the sequential algorithm fulfills our goal: It puts a significant number of sampling points (here indicated as $\text{▴}$ on the $\theta$-axis) at regions of interest, given by the two cusps in this specific example. Note that the maximum error is attained at the right cusp (located at $\theta = 110^\circ$) for both designs strategies, but the sequential design yields a substantially smaller error than the non-sequential design. The maximum error has decreased to 25% of the original value by an application of the sequential design strategy.

We have also compared the two design strategies for other subspaces of the DOF. These results are presented in Appendix A.2 and confirm the superiority of the sequential design approach. The sequential design strategy chooses more observations at locations where the function rises or descends steeply, or has a cusp-like structure. Thus, for data simulated from the RSW model the proposed sequential design strategy yields substantial improvements in the performance of the Kriging interpolator in contrast to a regular design grid consisting only of equally spaced sampling locations. In the following section we will continue these investigations and study the methodology for real data from atomistic simulations.

2.3. Sequential Designs in Atomistic Simulations

So far we showed that Kriging outperforms series interpolation and additionally that Kriging can be further improved by sequential sampling techniques. These observations were made on the basis of data which was generated from the RSW model. In this section we apply this methodology toward the sampling of grain boundary energy subspaces with atomistic simulations and demonstrate that, again, the sequential sampling technique outperforms a regular sampling technique with respect the accuracy.
of prediction and the ability of sampling at regions close to the (unknown) cusps.

2.3.1. Reference Data

As in the previous section, we want to evaluate the error which occurs due to an incomplete sampling of the parameter space. Instead of using the approximate, analytical RSW function for reference, we now generate a dense data set of energies from atomistic simulations, consisting of \( N_{\text{ref}} \) simulated grain boundaries plus additional special \( \Sigma \) grain boundaries (see Section Atomistic Data Generation), and interpolate between them. The atomistic reference data for the 1D subspaces of [100], [110], and [111] symmetric tilt grain boundaries (STGBs) is displayed in Figure 4, together with the linear interpolation as black lines. It shows that the interpolated function is much rougher than the idealized RSW function, which has been considered in Section 2.2. The rougher energy landscape makes the identification of the energy cusps more
challenging and requires a careful selection of the design points and interpolation parameters.

Note that Figures 2a,c seemingly show some outliers at angles close to 0° and \( \theta_{\text{max}} \), which are equal to 0 J m\(^{-2}\). These are data points for which the very small deviation from low-energy, stable structures leads to a relaxation of the atomic positions to these structures at the nearby angles. In a real application, this problem might be avoided by increasing the size of the atomistic model, but for our proof of principle we simply ignore these data points.

In the following, we first compare non-sequential and sequential design strategies for a fixed smoothing parameter \( \nu = 0.5 \).

Next we study the impact of the choice of this parameter on the quality of the Kriging estimator and its ability to detect cusps.

### 2.3.2. Effect of Sequential Design

In this subsection we focus on exploring the [100], [110], and [111] STGB subspaces by a Kriging interpolator with a fixed smoothing parameter \( \nu = 0.5 \). Corresponding results for \( \nu = 1.5, 2.5 \), and the maximum likelihood estimator \( \hat{\nu} \) updated in every step of the sequential algorithm can be found in Appendix C. We considered the three misorientation subspaces from 0° to \( \theta_{\text{max}} \), with \( \theta_{\text{max}} = 90° \) for the [100] subspace, \( \theta_{\text{max}} = 180° \) for the [110] subspace, and \( \theta_{\text{max}} = 120° \) for the [111] subspace. An initial design with \( N_{\text{init}} \) grain boundaries (16 for the [100], 31 for the [110], and 21 for the [111] subspace) and a sequential design with \( N_{\text{seq}} \) grain boundaries (15 for the [100] and [110] subspaces and 20 for the [111] subspace) were used for the exploration of the subspaces. Note that the new points were not restricted to a pre-defined set of misorientations, but chosen freely by the algorithm. However, for numerical reasons the minimum distance of misorientations was fixed to 0.05°. The resulting energies predicted by the Kriging interpolator are displayed in Figure 5 as black lines, and the values are given on the right y-axis. The left y-axis indicates the stage of the sequential design, which places the design points at specific misorientation angles.

We observe that most of the experiments in the sequential designs are conducted in neighborhoods of the cusps. This is remarkable, because the starting designs contain no misorientation angles corresponding to cusps. In particular, regions with large local energy gradients of the subspace are sampled in more detail by the sequential design. Compared to regular sampling, which does not guarantee the discovery of an unknown cusp of the subspace, this is an important advantage of the sequential approach. Note also that the sequential design does not access the cusps one by one in full detail, but visits them all repeatedly, step by step. It thus reduces the maximum error, which is usually located in the vicinity of the cusps. In this way, nearly every cusp of the STGB subspaces can be found within only a few iterations.

In Table 2 we display the maximum error of the Kriging interpolator based on the two sampling schemes. Compared to
the regular sampling technique the sequential approach leads to a decrease of the maximum error from 0.4560 to 0.0888 J m⁻² for the [100] subspace, from 0.2349 to 0.2161 J m⁻² for the [110] subspace, and from 0.2349 to 0.1699 J m⁻² for the [111] subspace.

An interesting question in this context is whether the error of the sequential design strategy always converges to a minimum value with increasing sample size. For large smoothness parameters $\nu$, an overfitting of the data, and hence no convergence, can be expected. For small $\nu$, however, adding data should always improve the fit. This is demonstrated in Figure 6, which shows the maximum error of the Kriging interpolator as a function of the sample size of the sequential design. For example, the black curve corresponds to the [100] subspace, where the sample size of the initial design is $N_{\text{init}} = 16$. The line shows the maximum error of prediction by the Kriging interpolator for the next 20 sequentially chosen points. We observe that at the beginning of the sequential phase the error decreases rapidly, because the sequential design allocates design points in neighborhoods of the cusps. Afterward, the error reaches a plateau and remains constant for several steps. This phenomenon can be explained by the choice of the maximum absolute deviation as an error measure, which only decreases substantially if design points are added close to the location where the maximum absolute deviation in Equation (1) is currently taken. Since the sequential design algorithm switches between different areas of the sampling domain in order to guarantee a sufficient exploration of the energy function over the whole sampling domain (see Figure 5), there might always be sequences of successive sequential design steps where the error will not change. In particular, in the example of the [100] subspace, no further improvement is obtained after 11 sequential experiments have been conducted. As consequence atomistic simulations could have been stopped already at this point for this specific experiment and the error measure under consideration. For the two other subspaces we observe a similar qualitative behavior.

![Figure 6](image)

### 2.3.3. Effect of the Smoothing Parameter $\nu$ in Atomistic Simulations

In this subsection we briefly investigate the sensitivity of the method with respect to the choice of the smoothing parameter $\nu$ in the Kriging interpolator, where we focus on the [110] subspace. Corresponding results for the two other subspaces show a similar picture and can be found in Appendix C. In Figure 7 we display the dynamics of the sequential design and the resulting Kriging interpolator for different values of $\nu = 0.5, 1.5, 2.5$, and $\nu$ calculated in an adaptive way (maximum likelihood estimation). We observe that the sequential design strategy chooses more points in neighborhoods of the cusps, independently of the choice of $\nu$, but the resolution of the sampling depends on the choice of $\nu$. While for $\nu = 0.5$ the sequential design allocates at least one point to every cusp of the subspace, the design points of the sequential design for $\nu = 2.5$ remain close to the three deepest cusps at 0°, 70.53° ($\Sigma 3$), and 180°. Sequential sampling of the [110] STGB subspace with a data adaptive choice of smoothing parameter (in the interval [0.5, 2.5]), which is updated in each iteration, showed the following behavior: the interpolation started with $\nu \approx 1.5$ for the initial design. Adding sequential points then caused an adjustment of $\nu$ toward smaller values, finally varying between 0.6 and 0.65. This means that an increasing number of sequentially chosen data points leads to a Kriging interpolator that is close to linear interpolation. In contrast to the data generated from the RSW model, where the choice $\nu = 2.5$ for the smoothing parameter yields the best performance of the Kriging interpolator, the results in this section demonstrate that for atomistic data small values of $\nu$ should be preferred. This is up to what should be expected as atomistic simulations produce much rougher curves than the idealized RSW model and as a consequence the parameter $\nu$ should be chosen much smaller for this type of data.

### 3. Conclusion

The numerical results for the RSW model demonstrate that the Kriging interpolator (with Matérn kernel) has the capacity to outperform the interpolator based on series estimation in a wide range of scenarios. The suggested sequential design algorithm can even further improve the performance of the Kriging interpolator due to its tendency to prefer observations at locations where the target energy function has higher fluctuations and is more difficult to estimate. Our results for atomistic simulations of 1D STGB subspaces demonstrate that sequential sampling techniques can lead to substantial improvements in the prediction of grain boundary energies. In particular, grain boundaries chosen by the sequential design are located close to cusps of the unknown energy functions, which are the most important areas...
to sample due to their high energy gradient. Moreover, the proposed sequential design strategy chooses grain boundaries for the atomistic simulation by jumping between different cusps and is therefore able to sample grain boundaries at multiple cusps simultaneously. Consequently, the sequential sampling technique can discover unknown cusps, which opens the opportunity to sample even more complex subspaces. The accuracy of the Kriging interpolator and the number of iterations required to find the cusps depend on the choice of a smoothness parameter but estimating this parameter via a maximum likelihood approach yields very satisfying results. Moreover, for atomistic simulations smaller values of the smoothness parameter close to 0.5 can be used as well. In this case the maximum error decreases with an increasing number of sequential steps and converges after a few sequential iterations.

Due to the different dimensions of the subspaces and datasets, our new sequential design strategy and Kriging interpolator cannot be directly compared with previous sampling methods.\cite{11,27} However, the obvious advantage is the ability of our method to identify the energy cusps. In ref. [13] a priori knowledge of their location is needed for a meaningful interpolation between the data, and also a satisfying training of an ANN\cite{27} requires the inclusion of at least some of the special grain boundaries in the training data. In contrast to these approaches we demonstrated in this paper that a Kriging interpolator combined with a sequential sampling technique is able to identify cusps of the subspaces automatically without any prior information concerning their location. Due to the convergence properties mentioned in the preceding paragraph one can keep the number of sample points that are additionally chosen by the sequential design to a minimum. However, our results also indicate that the initial design should not be too sparse in order to provide an initial guess of the energy function that is a good starting point for the sequential algorithm. This strategy will be advantageous for any application with strong, localized fluctuations in the values of the unknown function.

4. Experimental Section

The Read–Shockley–Wolf Model: The original Read–Shockley model\cite{26} relies on the fact that GB structures with small misorientation angles can be viewed as 2D arrays of lattice dislocations. In this case, elasticity theory predicts a logarithmic dependency of GB energy on the misorientation angle. Wolf\cite{35} extended the Read–Shockley model into a piecewise logarithmic function which spans the whole range of misorientation angles in the 1D subspace of STGBs, and can capture all cusps and maxima, by connecting the appropriate number of logarithmic segments. This fully empirical extension can be motivated by the existence of grain boundary dislocations with continuous and much smaller Burgers vectors than lattice dislocations.\cite{32}

Given specific misorientation angles $\beta = (\beta_0, \ldots, \beta_k)$ with $0 = \beta_0 < \beta_1 < \ldots < \beta_k = \theta_{\text{max}}$, which mark the beginning and ending of the individual logarithmic segments, and energy levels $\gamma = (\gamma_0, \ldots, \gamma_k)$ the RSW
function at the misorientation angle \( \theta \) is given by

\[
F_{\text{RSW}}(\theta, \beta, \gamma) = \sum_{i=1}^{k} 1_{\beta_i \in \beta}(\theta) \times \left[ \begin{array}{c}
1 - \cos(\gamma - \gamma_{i-1}) \cdot \sin(\beta - \beta_{i-1}) \\
+ 1_{\beta_i \in \beta}(\gamma - \gamma_{i-1}) \cdot \cos(\beta - \beta_{i-1})
\end{array} \right],
\]

(3)

where the function \( f_{\text{RSW}}(\theta) \) is defined by

\[
f_{\text{RSW}}(\theta) = \sin(\pi \theta / 2)(1 - a \log(\sin(\pi \theta / 2)))
\]

(4)

for some real valued parameter \( a \) (in the experiments, the authors exclusively consider \( a = 1 \)), and \( 1_{A} \) is the indicator function of the set \( A \), that is,

\[
1_{A}(\theta) = \begin{cases} 1, & \text{if } \theta \in A, \\ 0, & \text{if } \theta \notin A. \end{cases}
\]

The indicator function determines the orientation of the logarithmic segment, that is, it describes a non-descending \((\gamma_{i} - \gamma_{i-1} \geq 0)\) or descending \((\gamma_{i} - \gamma_{i-1} < 0)\) branch of the energy.

In this study, the authors define \( F_{\text{RSW}} \) for STGBs with misorientation axis [110] via \( k = 6 \) logarithmic segments between the following misorientation angles and energy offset values,

\[
\beta = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5)
\]

\[
= (0^\circ, 20.775^\circ, 50^\circ, 80.4199^\circ, 110^\circ, 150.49^\circ, 180^\circ),
\]

\[
y = (\gamma_0, \gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5)
\]

\[
= (0, 0.683, 0.285, 0.664, 0.043, 0.826, 0).
\]

**Trigonometric Series and Kriging Interpolation:** The authors were interested in the prediction of 1D sections of the whole grain boundary energy landscape from a data set of observations that should be as small as possible. This landscape can be parametrized by one single parameter, the misorientation angle. More precisely, the target quantity is the energy function

\[
y = f(\theta)
\]

(7)

that maps an angle \( \theta \in [0, \theta_{\text{max}}] \) to the corresponding grain boundary energy \( y \). Based on data \( D_N = \{(\theta_1, y_1), \ldots, (\theta_N, y_N)\} \) satisfying the functional relationship in Equation (7), the aim was to predict the energy at angles \( \theta \) where the corresponding energy had not been observed. Such a predictor is desirable since the direct computation of the energy both by physical or computer experiments is often exceptionally expensive.

In this section two approaches to tackle the prediction problem are proposed. Both methods have the desirable property that the resulting curves interpolate exactly through the observed data.

**Truncated Trigonometric Series Expansion:** The idea of the first approach was to model the function \( f \) in Equation (7) as a truncated trigonometric series. This was inspired by the use of spherical harmonics in materials science. For instance, the solid-liquid interface free energy was expanded in a series of hyperspherical harmonics for application in a phase field model of dendritic growth. \(^{10,11}\) In this context the function was used to represent the surface energy anisotropy, but it can also be used to model grain boundary energies. This was demonstrated by Banadaki and Patala,\(^{11}\) who fitted a series of symmetrized hyperspherical harmonics to investigate the dependence of interfacial energies on the boundary-plane inclination of grain boundaries. For the 1D energy functions of STGBs, the analogue of this representation of a 2D, spherical function by spherical harmonics, is the representation in terms of trigonometric basis functions.

\[
f(\theta) = \sum_{i=0}^{T} c_i \varphi_i(\theta),
\]

(8)

where the basis functions \( \varphi_i \) are defined by

\[
\varphi_i(\theta) = \frac{1}{\theta_{\text{max}}} 1_{[0, \theta_{\text{max}}]}(\theta),
\]

(9)

\[
\varphi_{2j-1}(\theta) = \sqrt{\frac{2}{\theta_{\text{max}}}} \cos \left( \frac{2(j-1)\theta}{\theta_{\text{max}}} \right),
\]

(9)

\[
\varphi_{2j}(\theta) = \sqrt{\frac{2}{\theta_{\text{max}}}} \sin \left( \frac{2j\theta}{\theta_{\text{max}}} \right), \quad j = 1, 2, \ldots
\]

(9)

The coefficients can be estimated from the data \( D_N \) by means of a least squares approach, that is, \( \hat{c} = (\hat{c}_0, \ldots, \hat{c}_T) \) is determined as the minimizer of the sum of squares

\[
\sum_{i=1}^{N} (y_i - \hat{f}(\theta_i))^2 = \sum_{i=1}^{N} \left( y_i - \sum_{j=0}^{T} c_j \varphi_j(\theta_i) \right)^2.
\]

(10)

The resulting function \( \hat{f}(\theta) = \sum_{i=0}^{T} c_i \varphi_i(\theta) \) is then used to predict energies at unobserved angles. If \( T = N - 1 \) (that is, the number of basis functions equals the number of observations), minimizing Equation (10) with respect to \( c_1, \ldots, c_T \) leads to a system of \( n \) linear equations for the parameters \( c_1, \ldots, c_T \) with a unique solution (at least if the sampling points form an equidistant grid of the interval \([0, \theta_{\text{max}}]\)). Then, the resulting function \( \hat{f} \) interpolates exactly through the given data as has been illustrated in Figure 1a above.

**Simple Kriging Interpolator:** The second approach that was considered is called the (simple) Kriging interpolator which is a popular technique in applied statistics for interpolation of scattered data. Originally proposed in the context of geostatistical applications\(^{16,30}\) it has also become a standard method in the area of computer experiments.\(^{28}\) In the case of noisy observations (that is, the observations in Equation (7) are perturbed by some random error) Kriging is also referred to as Gaussian process regression. The Kriging interpolator depends on the choice of a positive definite kernel function

\[
k(\theta, \theta'), \quad \theta, \theta' \in [0, \theta_{\text{max}}]
\]

(11)

(see Section 3.2 of ref. [22] for a rigorous definition of kernel functions). Using the data

\[
D_N = \{(\theta_1, y_1), \ldots, (\theta_N, y_N)\}
\]

(12)

the matrix

\[
K_{D_N} = (k(\theta_i, \theta_j))_{i,j=1}^{N} \in \mathbb{R}^{N \times N}
\]

(13)

and the vector \( k_{D_N} = (k(\theta, \theta_1), \ldots, k(\theta, \theta_N)) \in \mathbb{R}^{1 \times N} \). Here \( \theta \) is the point of interest where one wants to predict the energy. The matrix \( K_{D_N} \) is called Gramian matrix and is invertible for all combinations of kernel and sampling locations considered in this paper. The Kriging interpolator \( \hat{f}_{D_N} \) based on the set \( D_N \) of observations is finally given by the formula

\[
\hat{f}_{D_N}(\theta) = k_{D_N} K_{D_N}^{-1} y_T,
\]

(14)

where \( y = (y_1, \ldots, y_N) \in \mathbb{R}^{1 \times N} \) for \( y_i = f(\theta_i) \). An important advantage of the Kriging interpolator consists in the fact that it allows to quantify the
uncertainty of the prediction using a Bayesian interpretation. More precisely, modeling a random function as a centred Gaussian process with covariance kernel $k$, the Kriging interpolator is the posterior mean after observing the data $D_N$ \cite[Equation (2.19)]. In this case, the posterior variance (at location $\theta$) is given by

$$
\tau_{\theta_N}(\theta) = k(\theta, \theta) - k(\theta, \hat{y}_N) k_N^{-1} k_N^T D_N
$$

which does not depend on the observed responses $Y_1, \ldots, Y_N$. Roughly speaking, large values of $\tau_{\theta_N}$ indicate locations with a rather large uncertainty concerning the target function $f$ whereas small values are attained at locations where one has already a better understanding of $f$. In particular, the posterior variance $\tau_{\theta_N}$ vanishes at the points $\theta_j$, $\ldots$, $\theta_N$, where simulations have already been performed.

To specify the Kriging estimator, an appropriate kernel function was chosen. In this work, the Matérn class of kernels given through

$$
k(\theta, \theta') = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left( \sqrt{2\nu} |\theta - \theta'| / \theta_{\max} \right) ^\nu K_\nu \left( \sqrt{2\nu} |\theta - \theta'| / \theta_{\max} \right)
$$

was considered where $\nu > 0$ denotes a smoothness parameter, $\sigma > 0$ a variance parameter, and $K_\nu$ denotes a modified Bessel function. The hyperparameters $\nu$ and $\sigma$ had to be estimated (if we did not fix them a priori). The Matérn class contains a large class of interesting kernels as special cases (see Table 1 in ref. \cite{8}). For instance, in the case $\nu = 0.5$, the Kriging interpolator is nearly a piecewise linear function (provided that the scale parameter is not too small) and the corresponding kernel is referred to as the exponential kernel. Larger values of $\nu$ lead to smoother interpolators. A typical Kriging estimator for $\nu = 2.5$ is depicted in Figure 1b. In some machine learning applications, the (in terms of smoothness interpretable) parameter $\nu$ was fixed in advance (see, for instance, \cite{19}) for an application in wind field modeling with fixed choice $\nu = 2.5$, and only the remaining hyperparameter had to be estimated. In our study, both scenarios were considered where one either fixed the parameter $\nu$ in advance or one estimates $\nu$ from the data. As the method of choice for the estimation of hyperparameters the maximum likelihood approach that was also used in ref. \cite{14} was chosen, and which could briefly be summarized as follows: The construction of the Kriging estimator is based on the assumption that a priori the response vector $Y = (Y_1, \ldots, Y_N)$ follows a multivariate normal distribution with zero mean and covariance matrix $K_{\theta_N}$. Consequently, its probability density function is given by

$$
\varphi_Y(x) = \frac{1}{(2\pi)^{N/2} \det K_{\theta_N}} \exp \left( -\frac{1}{2} x K_{\theta_N}^{-1} x \right)
$$

for $x = (x_1, \ldots, x_N) \in \mathbb{R}^{N \times 1}$. The maximum likelihood approach then consists in choosing the unknown hyperparameters $(\sigma, \nu)$ as maximizers $(\hat{\sigma}, \hat{\nu})$ of the likelihood $\varphi_Y$ given the data $y = (y_1, \ldots, y_N) \in \mathbb{R}^{N \times 1}$, that is,

$$
(\hat{\sigma}, \hat{\nu}) \in \arg \max_{(\sigma, \nu)} \varphi_Y(y)
$$

In this context, the likelihood $\varphi_Y$ is understood as a function of the hyperparameters only whereas the data vector $y$ is fixed. Note that the dependence of the likelihood on the hyperparameters is given indirectly by means of the covariance $K_{\theta_N}$ \cite{13}, which directly depends on them through the function $k$ in Equation (16). Intuitively, the maximum likelihood approach chooses the hyperparameters such that the observed data are quite plausible.

**Sequential Designs for Kriging:** Given a limited budget concerning the number of possible experiments, the question of an appropriate design is of particular importance in material science simulations. For this purpose several strategies such as space-filling or Latin hypercube designs have been proposed in the statistical literature \cite{12,17,19,31}. Roughly speaking, non-sequential and sequential designs are distinguished. In the non-sequential case a design is fixed at the beginning and the experiments are conducted according to this design. In the sequential design only a part of the budgets used as an initial sample (as in the non-sequential case), and for the remaining observations the experimental conditions are sequentially updated depending on the output of the previous experiments. Both, sequential and non-sequential design strategies have their pros and cons (depending on the particular application), and the advantages of the Kriging interpolator based on a sequential design strategy for the prediction of grain boundary energies have been demonstrated in Sections 2.2 and 2.3.

The sequential design algorithm used in this paper is an adaption of an algorithm proposed in Kleijnen and van Beers\cite{14} to GB energy prediction. Given $N$ observations $D_N = \{(\theta_1, y_1), \ldots, (\theta_N, y_N)\}$ where $N \approx N_{\text{init}}$ the question how to find the $(N + 1)$th sampling location $\theta_{N+1}$ is answered. This procedure consists of the following steps:

1. **Selection of candidate points:** In a first step, $N_{\text{cand}}$ candidate points $\tilde{\theta}_1, \ldots, \tilde{\theta}_{N_{\text{cand}}}$ (among which $\theta_{N+1}$ will be selected) are determined via the following pseudo-code:

   - Set $D := D_N$
   - For $i = 1, \ldots, N_{\text{cand}}$:
     - Choose $\tilde{\theta}_i$ as the point in the experimental region $[0, \theta_{\max}]$, which maximizes the posterior variance $\tau_{\theta_N}(\tilde{\theta}_i)$ in Equation (15)
   - Set $D := D \cup \{\tilde{\theta}_1, \ldots, \tilde{\theta}_{N_{\text{cand}}}\}$

   Note that the definition of the posterior variance $\tau_{\theta_N}(\tilde{\theta})$ indeed depends only on the sampling locations but not on the response values. Therefore, it can be computed without performing any simulation-based measurements at all. The idea behind the above pseudo-code is to add successively sampling locations where the uncertainty (measured in the Bayesian interpretation by the value of the posterior variance) is maximal.

2. **Cross-validation:** In order to select the next sampling point $\theta_{N+1}$, the variance of the predicted output is estimated at each candidate input based only on Kriging computations which are computationally cheap in comparison with the expensive computations of the computer experiment of interest. For this purpose, Kriging estimators based on cross-validation are considered. Given the data $D_N$, $D_N^{(c)}$ denotes the data with the jth observation deleted, that is,

$$
D_N^{(c)} := \{(\theta_1, y_1), \ldots, (\theta_{j-1}, y_{j-1}), (\theta_{j+1}, y_{j+1}), \ldots, (\theta_N, y_N)\}
$$

Based on the sets $D_N$ and $D_N^{(c)}$ for $i = 1, \ldots, N$ the Kriging estimates are computed

$$
\hat{f}_{D_N}(\theta_i) \quad \text{for} \quad i = 1, \ldots, N_{\text{cand}},
$$

$$
\hat{f}_{D_N^{(c)}}(\theta_i) \quad \text{for} \quad j = 1, \ldots, N \text{ and } i = 1, \ldots, N_{\text{cand}}.
$$

at all the candidate points $\tilde{\theta}_i$, either based on the whole datasets $D_N$ or on the reduced datasets $D_N^{(c)}$.

3. **Computation of jackknife variance:** Using the Kriging estimates, jackknife’s pseudo-value for candidate $i$ is defined as

$$
\hat{\gamma}_i = N \hat{f}_{D_N}(\tilde{\theta}_i) - (N - 1) \hat{f}_{D_N^{(c)}}(\tilde{\theta}_i),
$$

for $j = 1, \ldots, N$ and $i = 1, \ldots, N_{\text{cand}}$. From the $\hat{\gamma}_i$ the jackknife variance for the candidate location $\tilde{\theta}_i$ is computed as

$$
\hat{\gamma}^2(\tilde{\theta}_i) = \frac{1}{N(N-1)} \sum_{j=1}^N (\hat{\gamma}_j - \hat{\gamma})^2,
$$

where $\hat{\gamma} = \frac{1}{N} \sum_{j=1}^N \hat{\gamma}_j.$
4) Selection of the next sampling location \( \theta_{n+1} \):

Finally, the winning candidate \( \theta_{n+1} \) among \( \tilde{\theta}_1, ..., \tilde{\theta}_{N_{\text{rand}}} \) is determined as

\[
\theta_{n+1} = \arg \max_{\tilde{\theta} \in \{\tilde{\theta}_1, ..., \tilde{\theta}_{N_{\text{rand}}}\}} \sum_{i} \left( \tilde{\theta} \right)_i^2
\]

Having determined \( \theta_{n+1} \), the corresponding response \( y_{n+1} \) is obtained from a (simulation) experiment. Afterwards, the sequential design algorithm can be used again to determine the next sampling location \( \theta_{n+2} \) and so on. The choice of \( N_{\text{rand}} \) permits the following heuristic: Taking \( N_{\text{rand}} = 1 \) leads to the, in the computer experiments literature well-known, algorithm of selecting the next design point among the maximizers of the posterior variance. In this special case, the next sampling point does not depend on the response values anymore and accumulation of sampling at region of interests (in our case, cusps) can not be achieved. For increasing \( N_{\text{rand}} \) more points close to already existing observations are allowed and the desired kind of accumulation becomes more likely. In all the simulations the authors put \( N_{\text{rand}} = 75 \).

Atomistic Simulation Method: To determine the grain boundary energy for a given set of DOF, a molecular statics simulation was performed using LAMMPS (Large-scale Atomic/Molecular Massively Parallel Energy) for a given set of DOF, a molecular statics simulation was performed using LAMMPS (Large-scale Atomic/Molecular Massively Parallel Simulator)\(^{[21]} \). Body centred cubic iron was chosen as an example, using the second embedded atom model (EAM) potential from \( [18] \) which predicted a lattice constant of Fe of \( 2.85\AA \). The grain boundary structures were generated by adapting the method introduced by Lee and Choi\(^{[15]} \) which allowed simulating non-periodic grain boundaries by using a spherical model. First, two spheres of atoms with a radius of \( r_o \) were created (in this work, \( r_o = 35a_{\text{Fe}} \)) and one of the spheres was rotated by the desired misorientation angle \( \phi \) around the rotation axis \( \beta \). Second, both grains were split in half-spheres and combined to create again a full sphere with a grain boundary with the desired normal vector \( \hat{n} \). Tied on the minimum grain boundary energy for this set of macroscopic degrees of freedom, the microscopic DOF had to be probed. For each set of macroscopic DOF, initial translations of 0 and 1.5 \( \AA \) perpendicular to the grain boundary plane were carried out, to probe the excess volume of the GB. Similarly, initial in-plane translations parallel to the grain boundary plane were carried out in steps of 0 to 5 \( \AA \) for the direction perpendicular to the tilt axis, and in steps of \( d_{\text{init}}/5 \) along the tilt axis, with \( d_{\text{init}} \) being the interplanar spacing along this direction, \( d_{\text{init}} = a/\sqrt{r^2 + k^2 + l^2} \) (\( hkl \)) are the Miller indices of the rotation axis. In addition to that a cutoff radius of either 0 or 0.25 \( \AA \) was used to delete overlapping atoms. In total, 144 different starting configurations of the microscopic DOF to relax the atomic positions and obtain the optimized structure for one set of macroscopic DOF of a grain boundary were used. After relaxation of all structures, the one with the minimum energy was kept in the database. Note that there is no guarantee that the minimum grain boundary energies are the absolute minimum energies for the respective set of macroscopic DOF, but an analysis of the relaxed trial structures showed that on average more than 25% of all initial sets of macroscopic DOF relax into the same state, which had the lowest energy of the set, or deviated by \( \leq 1\% \) from it. After relaxation of the atomic positions the grain boundary energy was calculated using only the atoms of an inner sphere with a radius of \( r_i \) (\( r_i = 20a_{\text{Fe}} \)), to avoid surface effects. Note that the difference between \( r_o \) and \( r_i \) had to be larger than the cut-off radius of the used interatomic potential plus the maximum applied parallel shift. The grain boundary energy was calculated according to the following equation

\[
E_{\text{GB}} = \frac{\sum_{n=1}^{N} E_{\text{pot}, n}}{\pi r_i^2},
\]

where \( N \) is the number of atoms in the inner sphere and \( E_{\text{pot}, n} \) the energy of the \( n \)-th atom in the inner sphere of all the initial macroscopic DOF, those which result in the minimum grain boundary energy for the given macroscopic DOF, were kept, and represented one datapoint in the energy landscape.

Table 3. Sampling parameters for each subspace, defined by rotation axis \( \hat{n} \):

| Subspace | \( \theta_{\text{max}}^\square \) | \( N_{\text{ref}} \) | \( N_{\text{total}} \) | \( N_{\text{init}} \) | \( N_{\text{seq}} \) | \( N_{\text{grid}} \) |
|----------|------------------|----------------|----------------|----------------|----------------|----------------|
| 100° STGBs | 90 | 97 | 46 | 16 | 20 | 1801 |
| 110° STGBs | 180 | 187 | 91 | 31 | 20 | 3601 |
| 111° STGBs | 120 | 125 | 61 | 21 | 20 | 2401 |

Atomistic Data Generation: The macroscopic DOF, in this case the misorientation angle, was varied between 0 and a maximum angle \( \theta_{\text{max}}^\square \) for each subspace, which was defined by the rotation axis: \( \theta_{\text{max}}^\square = 90° \) for [100] STGBs, 180° for [110] STGBs and 120° for [111] STGBs. A reference database with \( N_{\text{reg}} \) grain boundaries was generated, with a regular high throughput sampling technique and a spacing of \( \theta = 1° \), plus additional \( \Sigma \) grain boundaries, resulting in \( N_{\text{reg}} = 91 + 6, 181 + 8, \) and 121 + 6 data points, respectively. Information about the additional \( \Sigma \) grain boundaries can be found in Tables B1–c in Appendix B. The runs using the sequential sampling technique had \( N_{\text{init}} \) initial design points taken as equidistant subsets of the reference database. The points in the sequential design proposed in Section Sequential Designs for Kriging were chosen from an equidistant grid with

\[
N_{\text{grid}} = 2\theta_{\text{max}}^\square + 1
\]

points. Here the corresponding GB-energies were obtained from the reference database by linear interpolation.

Figure 6 shows that the maximum absolute error generally decreases with increasing number of sequential iterations, and it was decided to limit \( N_{\text{init}} \) to 20, to prove that the new technique is able to sample a subspace with as few grain boundaries as possible with a high accuracy. The information regarding the different sampling parameters is summarized in Table 3.

To analyze the influence of \( \nu \) towards the sampling of the subspace, first of all every subspace was sampled with the sequential sampling technique three times with different fixed \( \nu \) values each. The different \( \nu \) values chosen for the sampling were 0.5, 1.5, and 2.5. After that the sampling was repeated with \( \nu \) limited between 0.5 and 2.5 but estimated from the atomistic data using the MLE technique.

Appendix A: Further Results for The RSW Model

In this part of the Supporting Information we continue the investigation of Section 2.2 and provide results for five more test functions, which were considered in ref. [7] and are depicted in the left plots of Figures A1–A5, respectively. The functions \( F_4 \) and \( F_5 \) are examples of [100] and [111] STGBs. The RSW function \( F_{\text{RSW}} \) considered in Section 2.2 of the main part of the paper also belongs to this class, representing the [110] symmetric tilt grain boundary. On the other hand the functions \( F_4 \) and \( F_5 \) represent the [100], [110], and [111] twist grain boundaries. Note that not all functions satisfy the boundary condition \( f(0) = f(\theta_{\text{max}}) \) which is automatically satisfied by the trigonometric interpolator. For this reason we first transformed the data via \( y \rightarrow x ) = f(\theta_{\text{max}}) \) which is automatically satisfied by the trigonometric interpolator. For this reason we first transformed the data via \( y \rightarrow x = (a \theta + b) \) where the numbers \( a, b \) are chosen such that \( f(0) = f(\theta_{\text{max}}) \). Prediction was performed for the data transformed in this way, and afterward the predicted data were transformed back.

A.1. Comparison of Trigonometric Series and Kriging Interpolator

In analogy to Table 1 we state in Tables A1–A5 the maximum absolute error in Equation (1) for the remaining \( F_{\text{RSW}} \) functions for symmetrical...
tilt and twist grain boundaries with rotation axis \([100], [110], \) and \([111]\), and the interpolators under consideration. In all cases, different values for the parameter \(\nu\) including a data-adaptive choice \(\hat{\nu}\) by means of MLE were considered for the Kriging interpolator. These results confirm the findings from Section 2.2. For simulated data from the RSW model Kriging outperforms series interpolation for grain boundary energy prediction in all cases under consideration, where the value \(\nu = 2.5\) and MLE result in the best results for the Kriging estimator.

### A.2. Performance of The Sequential Design Algorithm

The following figures summarize the results concerning the sequential design algorithm for the \([100]\) and \([111]\) STGB subspaces \((F_1 \text{ and } F_3)\) and the \([100], [110], \) and \([111]\) twist grain boundaries subspaces \((F_4 \text{–} F_6)\). The illustration of our results is in complete analogy to the one used in Figure 2. The left plots show the respective dynamics of the sequential design algorithm for the best initial design leading to a final design of \(N_{\text{total}} = 65\) design points. As in Figure 2a, the \(\theta\)-value (angle) of the design points.
Figure A4. $F_5$: [110] twist grain boundaries subspace.

Figure A5. $F_6$: [111] twist grain boundaries subspace.

Table A1. $F_1$: [100] STGB subspace.

| N  | Trigonometric series | $\nu = 0.5$ | $\nu = 1.5$ | $\nu = 2.5$ | $\hat{\nu}$ via MLE |
|----|----------------------|-------------|-------------|-------------|------------------|
| 9  | 0.2163               | 0.1136      | 0.0949      | 0.0698      | 0.0925 ([$\nu$] = 1.5648) |
| 17 | 0.1206               | 0.0540      | 0.0434      | 0.0358      | 0.0365 ([$\nu$] = 1.9147) |
| 33 | 0.0690               | 0.0263      | 0.0206      | 0.0194      | 0.0198 ([$\nu$] = 1.8653) |
| 65 | 0.0384               | 0.0131      | 0.0100      | 0.0082      | 0.0086 ([$\nu$] = 1.7820) |

Table A2. $F_2$: [111] STGB subspace.

| N  | Trigonometric series | $\nu = 0.5$ | $\nu = 1.5$ | $\nu = 2.5$ | $\hat{\nu}$ via MLE |
|----|----------------------|-------------|-------------|-------------|------------------|
| 9  | 0.1020               | 0.0600      | 0.0519      | 0.0356      | 0.0357 ([$\nu$] = 2.4938) |
| 17 | 0.0573               | 0.0291      | 0.0240      | 0.0160      | 0.0160 ([$\nu$] = 2.4938) |
| 33 | 0.0318               | 0.0143      | 0.0114      | 0.0073      | 0.0073 ([$\nu$] = 2.4938) |
| 65 | 0.0175               | 0.0071      | 0.0055      | 0.0035      | 0.0035 ([$\nu$] = 2.4938) |

Table A3. $F_3$: [100] twist grain boundary subspace.

| N  | Trigonometric series | $\nu = 0.5$ | $\nu = 1.5$ | $\nu = 2.5$ | $\hat{\nu}$ via MLE |
|----|----------------------|-------------|-------------|-------------|------------------|
| 9  | 0.0414               | 0.0252      | 0.0229      | 0.0155      | 0.0156 ([$\nu$] = 2.4938) |
| 17 | 0.0228               | 0.0124      | 0.0105      | 0.0069      | 0.0069 ([$\nu$] = 2.4938) |
| 33 | 0.0125               | 0.0061      | 0.0049      | 0.0031      | 0.0031 ([$\nu$] = 2.4938) |
| 65 | 0.0068               | 0.0031      | 0.0024      | 0.0015      | 0.0015 ([$\nu$] = 2.4955) |

Table A4. $F_4$: [110] twist grain boundary subspace.

| N  | Trigonometric series | $\nu = 0.5$ | $\nu = 1.5$ | $\nu = 2.5$ | $\hat{\nu}$ via MLE |
|----|----------------------|-------------|-------------|-------------|------------------|
| 9  | 0.1461               | 0.1046      | 0.0871      | 0.0608      | 0.0655 ([$\nu$] = 2.2559) |
| 17 | 0.0812               | 0.0499      | 0.0403      | 0.0269      | 0.0270 ([$\nu$] = 2.4938) |
| 33 | 0.0451               | 0.0244      | 0.0191      | 0.0123      | 0.0154 ([$\nu$] = 1.9192) |
| 65 | 0.0252               | 0.0121      | 0.0093      | 0.0060      | 0.0074 ([$\nu$] = 1.9423) |

Table A5. $F_5$: [111] twist grain boundary subspace.

| N  | Trigonometric series | $\nu = 0.5$ | $\nu = 1.5$ | $\nu = 2.5$ | $\hat{\nu}$ via MLE |
|----|----------------------|-------------|-------------|-------------|------------------|
| 9  | 0.0581               | 0.0278      | 0.0238      | 0.0163      | 0.0164 ([$\nu$] = 2.4938) |
| 17 | 0.0333               | 0.0134      | 0.0110      | 0.0074      | 0.0074 ([$\nu$] = 2.4938) |
| 33 | 0.0186               | 0.0066      | 0.0052      | 0.0034      | 0.0034 ([$\nu$] = 2.4938) |
| 65 | 0.0103               | 0.0033      | 0.0025      | 0.0016      | 0.0016 ([$\nu$] = 2.4938) |

indicates the location in the interval [0, $\theta_{\text{max}}$], the $y$-value the stage in the sequential design algorithm (stage 0 corresponds to the initial design). The $F_{RSW}$-function from which data are evaluated is plotted in gray. The right plots state the maximum absolute error of the Kriging interpolator in combination with the sequential design algorithm in dependence of the initial design size $N_{\text{init}}$ and the total design size $N_{\text{total}}$.

The results confirm our finding from Section 2.2. In all cases under consideration the sequential design leads to a substantial improvement.
Table B1. Σ STGBs added to the reference data sets for the (a) [100], (b) [110], and (c) [111] STGB subspace, which is defined by the rotation axis \( \vec{\rho} \).

| \( \Sigma \) | 5  | 5  | 13 | 13 | 41 | 41 |
|---|---|---|---|---|---|---|
| \( \theta[^\circ] \) | 36.87 | 53.13 | 22.62 | 67.38 | 12.68 | 77.32 |

| \( \Sigma \) | 3  | 3  | 9  | 9  | 11 | 11 | 51 | 51 |
|---|---|---|---|---|---|---|---|---|
| \( \theta[^\circ] \) | 70.53 | 109.47 | 38.94 | 141.06 | 50.48 | 129.52 | 16.10 | 163.90 |

| \( \Sigma \) | 7  | 7  | 13 | 13 | 57 | 57 |
|---|---|---|---|---|---|---|
| \( \theta[^\circ] \) | 38.21 | 81.79 | 27.80 | 92.20 | 13.17 | 106.83 |

of the Kriging interpolator and allocates more design points to regions close to cusps.

Appendix B: Special STGBs of The Reference Data Sets

The data generation of the reference data sets is described in Section Atomistic Data Generation. In addition to the equally spaced design points of the regular sampling, the positions of the special, CSL based, grain boundaries were added. Their energies represent the deep cusps in the energy function. The orientation relationships of these grain boundaries are displayed in Table B1a for the [100] subspace, in Table B1b for the [110] subspace, and in Table B1c for the [111] subspace together with their \( \Sigma \) value.

Appendix C: Further Atomistic Simulation Results

In this section we will present additional results for the situation discussed in Section 2.3. The dynamics of the sequential design algorithm and the Kriging predictions for different subspaces are shown in Figure C1 ([100] STGB subspace), Figure C2 ([110] STGB subspace), and Figure C3 ([111] STGB subspace), respectively. As in the main part of the paper, different choices of the smoothness parameter including a data-adaptive choice via maximum likelihood estimation were considered. The positions of the design points (misorientation angle \( \theta \)) in the interval \([0, \theta_{\text{max}}] ([100]: \theta_{\text{max}} = 90[^\circ], [110]: \theta_{\text{max}} = 180[^\circ], [111]: \theta_{\text{max}} = 120[^\circ]) are indicated by a ▴. The left y-value indicates the stage in the sequential design algorithm (stage 0 corresponds to the initial design), the right y-value the grain boundary energy in \( Jm^{-2} \) calculated by the Kriging interpolator. As for the simulations described in Section 2.3 the sequential design places a significant number of points close to a priori unknown cusps. Again, we observe that the sequential design strategy chooses more points in neighborhoods of the cusps, independently of the choice of \( \nu \), but the resolution of the sampling depends on the choice of \( \nu \). Finally, we briefly discuss the data adaptive choice of the smoothing parameter by the maximum likelihood method. For the [100] and the [111] subspace the data adaptive choice of the smoothing parameter starts at \( \hat{\nu} = 0.85 \) and \( \hat{\nu} = 1.51 \), respectively. For both subspaces \( \hat{\nu} \) decreases to 0.5 during the sequential algorithm and stays close to this value after a few iterations. If the Kriging interpolator and the sequential design are used for exploring the [110] subspace, the algorithm starts with \( \hat{\nu} = 1.39 \), decreases during the iterations, but does not converge to 0.5. Instead it decreases toward 0.6 and varies between 0.6 and 0.65.

Figure C1. Dynamics of the sequential design algorithm (\( N_{\text{seq}} = 20 \) sequential design points) for exploring the [100] STGB subspace by a Kriging interpolator with different smoothing parameters. a) \( \nu = 0.5 \), b) \( \nu = 1.5 \), c) \( \nu = 2.5 \), and d) \( \hat{\nu} \) calculated by MLE.
Figure C2. Dynamics of the sequential design algorithm ($N_{\text{seq}} = 20$ sequential design points) for exploring the [110] STGB subspace by a Kriging interpolator with different smoothing parameters. a) $\nu = 0.5$, b) $\nu = 1.5$, c) $\nu = 2.5$, and d) $\hat{\nu}$ calculated by MLE.

Figure C3. Dynamics of the sequential design algorithm ($N_{\text{seq}} = 20$ sequential design points) for exploring the [111] STGB subspace by a Kriging interpolator with different smoothing parameters. a) $\nu = 0.5$, b) $\nu = 1.5$, c) $\nu = 2.5$, and d) $\hat{\nu}$ calculated by MLE.
Acknowledgements
This research has been supported by the German Research Foundation (DFG), project number 414750139.
Open access funding enabled and organized by Projekt DEAL.

Conflict of Interest
The authors declare no conflict of interest.

Author Contributions
Project planning and conceptualization, H.D. and R.J.; project supervision, H.D. and R.J.; implementation, numerical simulations, data processing, and visualization, M.K. and T.S.; interpretation, discussion, and paper planning, M.K., T.S., H.D., and R.J.; Writing - original draft, M.K. and T.S.; Writing - revising, M.K., T.S., H.D. and R.J. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement
The code for the RSW model, which can be used to reproduce all the results obtained for this model, is available at https://gitlab.com/kroll.martin/rsw-model. Data from atomistic simulations can be downloaded from https://git.noc.ruhr-uni-bochum.de/Schmalofski. Timo/atomistic-data-STGBs-1D/.

Keywords
atomistic simulations, data based materials science, grain boundaries, statistical methods

Received: December 19, 2021
Revised: March 27, 2022
Published online: April 14, 2022

[1] A. D. Banadaki, S. Patala, Comp. Mater. Sci. 2016, 112, 147.
[2] A. Banerjee, X. Wang, C. Fang, E. A. Wu, Y. S. Meng, Chem. Rev. 2020, 120, 6878.
[3] V. V. Bulatov, B. W. Reed, M. Kumar, Acta Mater. 2014, 65, 161.
[4] Z. C. Cordero, B. E. Knight, C. A. Schuh, Int. Mater. Rev. 2016, 61, 495.
[5] D. Cornford, I. T. Nabney, C. K. I. Williams, J. Nonparametr. Stat. 2002, 14, 43.
[6] N. A. C. Cressie, Statistics for Spatial Data, John Wiley & Sons, Inc., New York 1993.
[7] H. Dette, J. Gösmann, C. Greiff, R. Janisch, Acta Mater. 2017, 125, 145.
[8] P. Guttorp, T. Gneiting, Biometrika 2006, 93, 989.
[9] S. Z. Han, E.-A. Choi, S.-H. Lim, S. Kim, J. Lee, Prog. Mater. Sci. 2021, 117, 100720.
[10] T. Haxhiriai, A. Karma, F. Gonzales, M. Rappaz, Nat. Mater. 2006, 5, 660.
[11] J. J. Hoyt, M. Asta, A. Karma, Phys. Rev. Lett. 2001, 86, 5330.
[12] V. R. Joseph, Qual. Eng. 2016, 28, 28.
[13] H.-K. Kim, W.-S. Ko, H.-J. Lee, S. G. Kim, B.-J. Lee, Scripta Mater. 2011, 64, 1152.
[14] J. P. C. Kleijnne, W. C. M. van Beers, J. Oper. Res. Soc. 2004, 55, 876.
[15] B.-J. Lee, S.-H. Choi, Model. Simul. Mater. Sci. 2004, 12, 621.
[16] L. Li, P. Wang, Q. Shao, X. Huang, Chem. Soc. Rev. 2020, 49, 3072.
[17] M. D. McKay, R. J. Beckman, W. J. Conover, Technometrics 1979, 21, 239.
[18] M. I. Mendeleev, S. Han, D. J. Srolovitz, G. J. Ackland, D. Y. Sun, M. Asta, Philos. Mag. 2003, 83, 3977.
[19] M. D. Morris, T. J. Mitchell, J. Stat. Plan. Infer. 1995, 43, 381.
[20] M. M. Noack, G. S. Doerk, R. Li, J. K. Streit, R. A. Vaia, K. G. Yager, M. Fukuto, Sci. Rep. 2020, 10, 17663.
[21] D. L. Olmsted, S. M. Foiles, E. A. Holm, Acta Mater. 2009, 57, 3694.
[22] V. I. Paulsen, M. Raghupathi, An Introduction to The Theory of Reproducing Kernel Hilbert Spaces, Cambridge University Press, Cambridge 2016.
[23] S. Plimpton, J. Comput. Phys. 1995, 117, 1.
[24] C. E. Rasmussen, C. K. I. Williams, Gaussian Processes for Machine Learning, MIT Press, Cambridge, MA 2006.
[25] S. Ratanaphan, D. L. Olmsted, V. V. Bulatov, E. A. Holm, A. D. Rollett, G. S. Rohrer, Acta Mater. 2015, 88, 346.
[26] W. T. Read, W. Shockley, Phys. Rev. 1950, 78, 275.
[27] S. E. Restrepo, S. T. Giraldo, B. J. Thijsse, Comp. Mater. Sci. 2014, 86, 170.
[28] T. J. Santner, B. J. Williams, W. I. Notz, The Design and Analysis of Computer Experiments, Springer, New York 2018.
[29] I. P. Semenova, R. Z. Valiev, A. M. Smyslov, M. V. Pesin, T. G. Langdon, Adv. Eng. Mater. 2021, 23, 2100145.
[30] M. L. Stein, Interpolation of Spatial Data, Springer, New York 1999.
[31] E. Stinstra, D. den Hertog, P. Stehouwer, A. Vestjens, Technometrics 2003, 45, 340.
[32] A. P. Sutton, R. W. Baluffi, Interfaces in Crystalline Materials, Oxford Science Publications, Oxford 1995.
[33] M. A. Tschopp, K. N. Solanki, F. Gao, X. Sun, M. A. Khaleel, M. Horstmeier, Phys. Rev. B 2012, 85, 064108.
[34] M. A. Tschopp, S. P. Coleman, D. L. McDowell, Int. Mater. 2015, 4, 176.
[35] D. Wolf, Scripta Metall. 1989, 23, 1713.
[36] L. Yang, C. Lai, S. Li, Comp. Mater. Sci. 2019, 162, 268.