SRSD: Rethinking Datasets of Symbolic Regression for Scientific Discovery

Yoshitomo Matsubara*
Amazon Alexa AI
yoshitom@uci.edu

Naoya Chiba, Ryo Igarashi & Yoshitaka Ushiku
OMRON SINIC X Corporation
{naoya.chiba,ryo.igarashi,yoshitaka.ushiku}@sinicx.com

Abstract

Symbolic Regression (SR) is a task of recovering mathematical expressions from given data and has been attracting attention from the research community to discuss its potential for scientific discovery. However, the community lacks datasets of symbolic regression for scientific discovery (SRSD) to discuss the potential of SR. To address the critical issue, we revisit datasets of SRSD to discuss the potential of symbolic regression for scientific discovery. Focused on a set of formulas used in the existing datasets based on Feynman Lectures on Physics, we recreate 120 datasets to discuss the performance of SRSD. For each of the 120 SRSD datasets, we carefully review the properties of the formula and its variables to design reasonably realistic sampling ranges of values so that our new SRSD datasets can be used for evaluating the potential of SRSD such as whether or not an SR method can (re)discover physical laws from such datasets. We conduct experiments on our new SRSD datasets using five state-of-the-art SR methods in SRBench, and the results show that the new SRSD datasets are more challenging than the original ones. Our datasets [1,2,3] and code repository [4] are publicly available.

1 Introduction

Recent advances in machine learning (ML), especially deep learning (DL), have led to the proposal of many methods that can reproduce the given data and make appropriate inferences on new inputs. Such methods are, however, often black-box, which makes it difficult for humans to understand how they made predictions for given inputs. This property will be more critical especially when non-ML experts apply ML to problems in their research domains such as physics and chemistry.

Symbolic regression (SR) is the task of producing a mathematical expression (symbolic expression) that fits a given dataset. SR has been studied in the genetic programming (GP) community [1–6], and DL-based SR has been attracting more attention from the ML/DL community [7–12]. Because of its interpretability, various scientific communities apply SR to advance research in their scientific fields *e.g.*, Physics [13–19], Applied Mechanics [20], Climatology [21], Materials [22–25], and Chemistry [26].

*This work was mainly done while the first author was a research intern at OMRON SINIC X Corporation.

[1] https://huggingface.co/datasets/yoshitomo-matsubara/srsd-feynman_easy

[2] https://huggingface.co/datasets/yoshitomo-matsubara/srsd-feynman_medium

[3] https://huggingface.co/datasets/yoshitomo-matsubara/srsd-feynman_hard

[4] https://github.com/omron-sinicx/srsd-benchmark

NeurIPS 2022 AI for Science Workshop.

Given that SR has been studied in various communities, La Cava et al. [11] propose SRBench, a unified benchmark framework for symbolic regression methods. In the benchmark study, they combine the Feynman Symbolic Regression Database (FSRD) [14] and the ODE-Strogatz repository [27] to compare a number of SR methods, using a large-scale heterogeneous computing cluster.[5]

To discuss the potential of symbolic regression for scientific discovery (SRSD), there still remain some issues to be addressed: oversimplified datasets and lack of evaluation metric towards SRSD. For symbolic regression tasks, existing datasets consist of values sampled from limited domains such as in range of 1 to 5, and there are no large-scale datasets with reasonably realistic values that capture the properties of the formula and its variables. Thus, it is difficult to discuss the potential of symbolic regression for scientific discovery with such existing datasets. For instance, the FSRD consists of 120 formulas selected mostly from Feynman Lectures Series[6] [32–34] and are core benchmark datasets used in SRBench [11]. While the formulas indicate physical laws, variables and constants used in each dataset have no physical meanings and sampling processes are oversimplified since the datasets in the benchmark study are not designed to discover the physical laws from the observed data in the real world. (See Section 3.1.)

To address these issues, we propose new SRSD datasets and conduct benchmark experiments using representative SR methods. We carefully review and design annotation policies for the new datasets, considering the properties of the physics formulas. Using the proposed SRSD datasets, we perform benchmark experiments with a set of symbolic regression baselines and find that even state of the art symbolic regression methods still need improvements to be used for scientific discovery.

2 Related Studies

In this section, we briefly introduce related studies focused on 1) symbolic regression for scientific discovery and 2) symbolic regression dataset and evaluation.

2.1 SRSD: Symbolic Regression for Scientific Discovery

A pioneer study on symbolic regression for scientific discovery is conducted by Schmidt and Lipson [35], who propose a data-driven scientific discovery method. They collect data from standard experimental systems like those used in undergrad physics education: an air-track oscillator and a double pendulum. Their proposed algorithm detects different types of laws from the data such as position manifolds, energy laws, and equations of motion and sum of forces laws.

Following the study, data-driven scientific discovery has been attracting attention from research communities and been applied to various domains such as Physics [13–19], Applied Mechanics [20], Climatology [21], Materials [22–25], and Chemistry [26].

These studies leverage symbolic regression in different fields. While general symbolic regression tasks use synthetic datasets with limited sampling domains for benchmarks, many of the SRSD studies collect data from the real world and discuss how we could leverage symbolic regression toward scientific discovery.

While SRSD tasks share the same input-output interface with general symbolic regression (SR) tasks (*i.e.*, input: dataset, output: symbolic expression), we differentiate SRSD tasks in this study from general SR tasks by whether or not the datasets including true symbolic expressions are created with reasonably realistic assumptions for scientific discovery such as meaning of true symbolic expressions (whether or not they have physical meanings) and sampling domains for input variables.

2.2 Dataset and Evaluation

For symbolic regression methods, there exist several benchmark datasets and empirical studies. The Feynman Symbolic Regression Database [14] is one of the largest symbolic regression datasets, which consists of 100 physics-inspired equations based on Feynman Lectures on Physics [32–34]. By randomly sampling from small ranges of value, they generate the corresponding tabular datasets for the 100 equations. Inspired by [1, 2, 4], Uy et al. [5] suggest 10 different real-valued symbolic

[5]Hosts with 24-28 core Intel(R) Xeon(R) CPU E5-2690 v4 @ 2.60GHz processors and 250 GB of RAM [11]

[6]Udrescu and Tegmark [14] extract 20 of the 120 equations as “bonus” from other seminal books [28–31].

regression problems (functions) and create the corresponding dataset (*a.k.a.* Nguyen dataset). The suggested functions consist of either 1 or 2 variables *e.g.*, $f(x) = x^6 + x^5 + x^4 + x^3 + x^2 + x$ and $f(x, y) = \sin(x) + \sin(y^2)$. They generate each dataset by randomly sampling 20 - 100 data points.

La Cava et al. [11] design a symbolic regression benchmark, named SRBench, and conduct a comprehensive benchmark experiment, using existing symbolic regression datasets such as the Feynman Symbolic Regression Database [14] and ODE-Strogatz repository [36]. In SRBench, symbolic regression methods are assessed by 1) an error metric based on squared error between target and estimated values, and 2) solution rate that shows a percentage of the estimated symbolic regression models that match the true models (equations).

However, these datasets and evaluations are not necessarily designed to discuss symbolic regression for scientific discovery. In Sections 3.1 and 4.1, we will further describe potential issues in such existing studies.

3 Datasets

In this section, we summarize issues we found in the existing symbolic regression datasets, and then propose new datasets to address them towards symbolic regression for scientific discovery (SRSD).

3.1 Issues in Existing Datasets

As introduced in Section 2.2, there are many symbolic regression datasets. However, we consider that novel datasets are required to discuss SRSD for the following reasons:

1. **No physical meaning:** Many of the existing symbolic regression datasets [1, 2, 4, 5] are not necessarily physics-inspired, but instead randomly generated *e.g.*, $f(x) = \log(x)$, $f(x, y) = xy + \sin((x-1)(y-1))$. To discuss the potential of symbolic regression for scientific discovery, we need to further elaborate datasets, considering how we would leverage symbolic regression in practice.
2. **Oversimplified sampling process:** While some of the datasets are physics-inspired such as the Feynman Symbolic Regression Database (FSRD) [14] and ODE-Strogatz repository [36], their sampling strategies are very simplified. Specifically, the strategies do not distinguish between constants and variables *e.g.*, speed of light[7] is treated as a variable and randomly sampled in range of 1 to 5. Besides, most of the sampling domains are far from values we could observe in the real world *e.g*, II.4.23 in Table S1 (the vacuum permittivity values are sampled from range of 1 to 5). When sampled ranges of the distributions are narrow, we cannot distinguish Lorentz transformation from Galilean transformation *e.g.* I.15.10 and I.16.6 in Table S3, I.48.2 in Table S5, I.15.3t, I.15.3x, and I.34.14 in Table S7, or the black body radiation can be misestimated to Stephan-Boltzmann law or the Wien displacement law *e.g.* I.41.16 in Table S8.
3. **Duplicate equations:** Due to the two issues above, many of the equations in existing datasets turn out to be duplicate. *e.g.*, as shown in Table 1, $F = \mu N_n$ (I.12.1) and $F = q_2 E$ (I.12.5) in the *original* Feynman Symbolic Regression Database are considered identical since both the equations are multiplicative and consists of two variables, and their sampling domains (Distributions in Table 1) are exactly the same. For instance, approximately 25% of the symbolic regression problems in the *original* FSRD have 1 - 5 duplicates in that regard.
4. **Incorrect/Inappropriate formulas:** The Feynman Symbolic Regression Database [14] treat every variables as float whereas they should be integer to be physically meaningful. For example, the number of phase difference in Bragg's law should be integer but sampled as real number (I.30.5 in Table S1). Furthermore, they don't even give special treatment of angle variables (I.18.12, I.18.16, and I.26.2 in Table 1). Physically some variables can be negative whereas the *original* Feynman Symbolic Regression Database [14] only samples positive values (*e.g.* I.8.14 and I.11.19 in Table S3). We also avoid using *arcsin*/*arccos* in the equations since the use of *arcsin*/*arccos* in the Feynman Symbolic Regression Database [14] just to obtain angle variable is not experimentally meaningful (I.26.2 in Table 1, I.30.5 in Table S1, and B10 in Table S11). Equations using *arcsin* and *arccos* in the original annotation are I.26.2 (Snell's law), I.30.5

[7]We treat speed of light as a constant ($2.998 \times 10^8 \mathrm{m/s}$) in this study.

(Bragg's law), and B10 (Relativistic aberration). These are all describing physical phenomena related to two angles, and it is an unnatural deformation to describe only one of them with an inverse function. Additionally, inverse function use implicitly limits the range of angles, but there is no such limitation in the actual physical phenomena.

3.2 Proposed SRSD Datasets

We address the issues in existing datasets above by proposing new SRSD datasets based on the equations used in the FSRD [14]. *i.e.*, Section 3.1 summarizes the differences between the FSRD and our SRSD datasets. Our annotation policy is carefully designed to simulate typical physics experiments so that the SRSD datasets can engage studies on symbolic regression for scientific discovery in the research community.

3.2.1 Annotation policy

We thoroughly revised the sampling range for each variable from the annotations in the FSRD [14]. First, we reviewed the properties of each variable and treated physical constants (*e.g.*, light speed, gravitational constant) as constants while such constants are treated as variables in the original FSRD datasets. Next, variable ranges were defined to correspond to each typical physics experiment to confirm the physical phenomenon for each equation. We also used [37] as a reference. In cases where a specific experiment is difficult to be assumed, ranges were set within which the corresponding physical phenomenon can be seen. Generally, the ranges are set to be sampled on log scales within their orders as 10^2 in order to take both large and small changes in value as the order changes. Variables such as angles, for which a linear distribution is expected are set to be sampled uniformly. In addition, variables that take a specific sign were set to be sampled within that range. Tables 1 and S1 – S11 show the detailed comparisons between the original FSRD and our proposed SRSD datasets.

3.2.2 Complexity-aware Dataset Categories

While the proposed datasets consist of 120 different problems, there will be non-trivial training cost required to train a symbolic regression model for all the problems individually [11] *i.e.*, there will be 120 separate training sessions to assess the symbolic regression approach. To allow more flexibility in assessing symbolic regression models for scientific discovery, we define three clusters of the proposed datasets based on their complexity: *Easy*, *Medium*, and *Hard* sets, which consist of 30, 40, and 50 different problems respectively.

We define the complexity of problem, using the number of operations to represent the true equation tree and range of the sampling domains. The former measures how many mathematical operations compose the true equation such as *add*, *mul*, *pow*, *exp*, and *log* operations. The latter considers magnitude of sampling distributions (*Distributions* column in Tables 1 and S1 – S11) and increases the complexity when sampling values from wide range of distributions. We define the domain range as follows:

$$f_{\text{range}}\left(\mathcal{S}\right) = \left|\log_{10}\left|\max_{s \in \mathcal{S}} s - \min_{s \in \mathcal{S}} s\right|\right|, \tag{1}$$

where $\mathcal{S}$ indicates a set of sampling domains (*distributions*) for a given symbolic regression problem.

As we will show in Section 5.3, these clusters represent problem difficulties at high level. For instance, these subsets will help the research community to shortly tune and/or perform sanity-check new approaches on the *Easy* set (30 problems) instead of using the whole datasets (120 problems). Figure 1 shows the three different distribution maps of our proposed datasets. *Easy*, *Medium*, and *Hard* sets consist of 30, 40, and 50 individual symbolic regression problems, respectively.

4 Benchmark

Besides the conventional metrics, we propose a new metric to discuss the performance of symbolic regression for scientific discovery in Section 4.1. Following the set of metrics, we design an evaluation framework of symbolic regression for scientific discovery.

Table 1: Easy set of our proposed datasets (part 1). C: Constant, V: Variable, F: Float, I: Integer, P: Positive, N: Negative, NN: Non-Negative, $\mathcal{U}$: Uniform distribution, $\mathcal{U}_{\log}$: Log-Uniform distribution.

Eq. ID	Formula	Symbols		Properties		Distributions	
				Original	Ours	Original	Ours
I.12.1	$F = \mu N_{\text{n}}$	F	Force of friction	V, F	V, F, P	N/A	N/A
		μ	Coefficient of friction	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-2}, 10^{0})$
		N_{n}	Normal force	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-2}, 10^{0})$
I.12.4	$E = \frac{q_1}{4\pi\epsilon r^2}$	E	Magnitude of electric field	V, F	V, F	N/A	N/A
		q_1	Electric charge	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$
		r	Distance	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-2}, 10^{0})$
		ϵ	Vacuum permittivity	V, F	C, F, P	$\mathcal{U}(1, 5)$	8.854×10^{-12}
I.12.5	$F = q_2 E$	F	Force	V, F	V, F	N/A	N/A
		q_2	Electric charge	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$
		E	Electric field	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{1}, 10^{3})$
I.14.3	$U = mgz$	U	Potential energy	V, F	V, F, P	N/A	N/A
		m	Mass	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-2}, 10^{0})$
		g	Gravitational acceleration	V, F	C, F, P	$\mathcal{U}(1, 5)$	9.807×10^{0}
		z	Height	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-2}, 10^{0})$
I.14.4	$U = \frac{k_{\text{spring}} x^2}{2}$	U	Elastic energy	V, F	V, F, P	N/A	N/A
		k_{spring}	Spring constant	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{2}, 10^{4})$
		x	Position	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-2}, 10^{0})$
I.18.12	$\tau = rF \sin\theta$	τ	Torque	V, F	V, F	N/A	N/A
		r	Distance	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		F	Force	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		θ	Angle	V, F	V, F, NN	$\mathcal{U}(0, 5)$	$\mathcal{U}(0, 2\pi)$
I.18.16	$L = mrv \sin\theta$	L	Angular momentum	V, F	V, F	N/A	N/A
		m	Mass	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		r	Distance	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		v	Velocity	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		θ	Angle	V, F	V, F, NN	$\mathcal{U}(1, 5)$	$\mathcal{U}(0, 2\pi)$
I.25.13	$V = \frac{q}{C}$	V	Voltage	V, F	V, F	N/A	N/A
		q	Electric charge	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-5}, 10^{-3})$
		C	Electrostatic Capacitance	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-5}, 10^{-3})$
I.26.2	$n = \frac{\sin\theta_1}{\sin\theta_2}$	n	Relative refractive index	V, F	V, F, P	$\mathcal{U}(0, 1)$	N/A
		θ_1	Refraction angle 1	V, F	V, F	N/A	$\mathcal{U}(0, \frac{\pi}{2})$
		θ_2	Refraction angle 2	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}(0, \frac{\pi}{2})$
I.27.6	$f = \frac{1}{\frac{1}{d_1} + \frac{n}{d_2}}$	f	Focal length	V, F	V, F	N/A	N/A
		d_1	Distance	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$
		n	Refractive index	V, F	V, F, P,	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		d_2	Distance	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$

4.1 Metrics

In general, it would be difficult to define "accuracy" of symbolic regression models since we will compare its estimated equation to the ground truth equation and need criteria to determine whether or not it is "correct". La Cava et al. [11] suggested a reasonable definition of symbolic solution, which is designed to capture symbolic regression models that differ from the true model by a constant or scalar. Using R^2 score (Eq. (2)), they also defined as accuracy the percentage of symbolic regression problems that a model meets $R^2 > \tau$, where τ is a threshold *e.g.*, $\tau = 0.999$ in [11].

$$R^2 = \frac{\sum_j^N \left(f_{\text{pred}}\left(X_j\right) - f_{\text{true}}\left(X_j\right)\right)^2}{\sum_k^N \left(f_{\text{true}}\left(X_k\right) - \bar{y}\right)^2}, \tag{2}$$

where N indicates the number of test samples (*i.e.*, the number of rows in the test dataset), and $\bar{y}$ is a mean of target outputs produced by f_{true}. f_{pred} and f_{true} are a trained SR model and a true model, respectively.

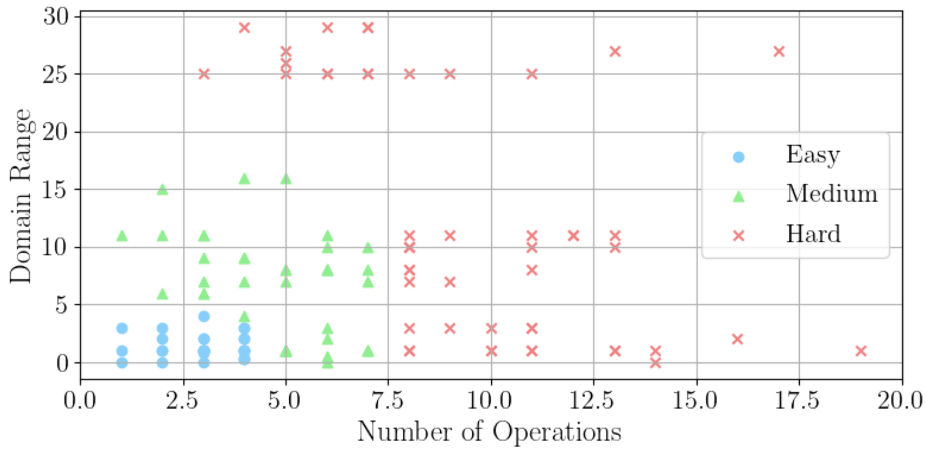

Figure 1: Distribution map of our proposed datasets based on three different subsets with respect to our complexity metrics. Data points at top right/bottom left indicate more/less complex problems.

4.2 Model Selection

For real datasets (assuming observed datasets), only tabular data are available for training and validation. (In practice, a test dataset does not include the true equation). For benchmark purposes, true equations are provided as test data besides test tabular data.

For each problem, we use the validation tabular dataset and choose the best trained SR model f^*_{pred} from $\mathcal{F}$, a set of the trained models by a given method respect to Eq. (3)

$$f^*_{\text{pred}} = \underset{f_{\text{pred}} \in \mathcal{F}}{\arg\min} \frac{1}{n} \sum_{i=1}^{n} \left| \frac{f_{\text{pred}}(X_i) - f_{\text{true}}(X_i)}{f_{\text{true}}(X_i)} \right|^2, \tag{3}$$

where X_i indicates the i-th row of the validation tabular dataset X.

We use the geometrical distance between predicted values against a validation tabular dataset to choose the best model obtained through hyperparameter tuning. Using the best model per method, we compute R^2 score to assess the method.

5 Experiments

5.1 Baseline Methods

For baselines, we use the five best symbolic regression methods in SRBench [11]. Specifically, we choose gplearn [3], AFP [38], AFP-FE [35], AI Feynman [15], and DSR [7], referring to the rankings of solution rate for the FSRD datasets in their study.

1. **gplearn** [3]: a genetic programming based symbolic regression method published as a Python package `gplearn`.
2. **AFP** [38]: Age-fitness pareto optimization.
3. **AFP-FE** [35]: AFP optimization with fitness estimates.
4. **AI Feynman** [15]: an iterative approach to generate symbolic regression to seek to fit data to formulas that are Pareto-optimal.
5. **DSR** [7]: reinforcement learning based deep symbolic regression.

For the details of the baseline models, we refer readers to the corresponding papers [3, 7, 15, 35, 38].

5.2 Runtime Constraints

The implementations of the baseline methods in Section 5.1 do not use any GPUs. We run 600 high performance computing (HPC) jobs in total, using “C.small” and “C.large” computing nodes, which

Table 2: Baseline results: accuracy ($R^2 > 0.999$) defined by La Cava et al. [11].

SRSD Datasets \ Method	gplearn	AFP	AFP-FE	AI Feynman	DSR
Easy set (30 problems)	6.67%	20.0%	23.3%	33.3%	**60.0%**
Medium set (40 problems)	7.50%	5.00%	5.00%	5.00%	**45.0%**
Hard set (50 problems)	2.00%	4.00%	4.00%	8.00%	**30.0%**

Table 3: Baseline results: solution rate defined by La Cava et al. [11].

SRSD Datasets \ Method	gplearn	AFP	AFP-FE	AI Feynman	DSR
Easy set (30 problems)	6.67%	20.0%	20.0%	30.0%	**43.3%**
Medium set (40 problems)	2.50%	2.50%	2.50%	2.50%	**10.0%**
Hard set (50 problems)	0.00%	0.00%	0.00%	**2.00%**	**2.00%**

Table 4: Solution rates of common baselines for FSRD and SRSD (Easy, Medium, Hard) datasets.

Dataset \ Method	gplearn	AFP	AFP-FE	AI Feynman	DSR
FSRD [14]	15.5%	20.48%	26.23%	**52.65%**	19.71%
SRSD (Ours)	1.67%	5.83%	5.83%	9.17%	**15.0%**

have 5 - 20 assigned physical CPU cores, 30 - 120 GB RAM, and 720 GB local storage available in AI Bridging Cloud Infrastructure (ABCI).[8] Due to the properties of our HPC resource, we have some runtime constraints:

1. Since each HPC job is designed to run for up to 24 hours due to the limited resource, we run a job with a pair of a target tabular dataset and a symbolic regression method.
2. Given a pair of a dataset and a method, each of our HPC jobs runs up to 100 separate training sessions with different hyperparameter values.

5.3 Results

In this section, we discuss the experimental results of our baseline methods, using the proposed SRSD datasets. Tables 2 and 3 show the performance of the symbolic regression baseline methods in terms of R^2-driven accuracy ($R^2 > 0.999$) and solution rate respectively, and both the metrics are used in SRBench [11]. According to the metrics, DSR significantly outperforms all the other baselines we considered. The DSR results also indicate difficulty levels of the three categories of our SRSD datasets, which looks aligned with our complexity-aware dataset categorization (Section 3.2.2).

Table 4 compares the solution rates of the five common baselines for the FSRD and our SRSD datasets. We can confirm that the overall solution rates for our SRSD are significantly degraded compared to those for the FSRD reported in SRBench [11]. The results indicate that our SRSD datasets are more challenging than the FSRD in terms of solution rate.

6 Limitations and Discussion

6.1 Implicit Functions

Symbolic regression generally has a limitation in inferring implicit functions, as the model infers a trivial constant function if there are no restrictions on variables. For example, $f(x, y) = 0$ is inferred as $0 = 0 \; \forall x, y$. This problem can be solved by applying the constraint that an inferred function should depend on at least two variables *e.g.*, inferring $f(x, y) = 0$ with $\frac{\partial f}{\partial x} \neq 0$ and $\frac{\partial f}{\partial y} \neq 0$, or by converting the function to an explicit form *e.g.*, $y = g(x)$. We converted some functions in the datasets into explicit forms and avoided the inverse trigonometric functions as described in Section 3.1.

[8] `https://abci.ai/en/how_to_use/tariffs.html`

6.2 Dummy Variables and Noise Injection

When applying machine learning to real-world problems, it is often true that 1) not all the observed features (variables in symbolic regression) are necessary to solve the problems, and 2) the observed values contain some noise. While we follow [11] and show experimental results for our SRSD datasets with noise-injected target variables in the supplementary material, these aspects are not thoroughly discussed in this study, such discussions can be a separate paper built on this work and further engage studies of symbolic regression for scientific discovery.

6.3 Interpretability Evaluation

Though symbolic regression methods are popular for intrpretability in their behaviors/outputs, there is a lack of appropriate metrics to evaluate these methods, taking into account the property. One of the most common approaches would be to measure the prediction error or correlation between the predicted values and the target values in the test data, as in standard regression problems. However, low prediction errors could be achieved even by complex models that differ from the original law.

Some studies [11, 15] use complexity of the predicted expression as an evaluation metric (the simpler the better). However, it is based on a big assumption that a simpler expression may be more likely to be a hidden law in the data (scientific discovery such as physics law), which may not be true for SRSD. SRBench [11] present the percentage of agreement between the target and the estimated equations, using solution rate they defined. But in such cases, both 1) equations that do not match at all and 2) that differ by only one term[9] are equally treated as incorrect. As a result, it is considered as a coarse-resolution evaluation method for accuracy in SRSD, which still needs more discussion towards real-world SRSD applications.

7 Conclusion

In this work, we pointed out issues of existing datasets and benchmarks of symbolic regression for scientific discovery (SRSD). To address the issues, we proposed 120 new SRSD datasets based on a set of physics formulas in FSRD [14] and conducted benchmark experiments using the proposed SRSD datasets. The results show the new SRSD datasets are significantly more challenging than the original FSRD datasets in terms of solution rate. We also discussed the limitations in this study and pointed out lack of evaluation metrics suitable for SRSD. Matsubara et al. [39] further discuss the issue and propose a new evaluation metric for SRSD, performing benchmark experiments with an additional baseline. To encourage the studies of SRSD, we publish our datasets[1, 2, 3] and code repository[4] with MIT License.

Acknowledgments

We thank the anonymous reviewers for their comments. In this study, we used the computing resource of AI Bridging Cloud Infrastructure (ABCI) provided by National Institute of Advanced Industrial Science and Technology (AIST). This work is supported by JST-Mirai Program Grant Number JPMJMI21G2, Japan.

References

[1] Nguyen Xuan Hoai, Robert Ian McKay, D Essam, and R Chau. Solving the Symbolic Regression Problem with Tree Adjunct Grammar Guided Genetic Programming: The Comparative Results. In *Proceedings of the 2002 Congress on Evolutionary Computation. CEC'02 (Cat. No. 02TH8600)*, volume 2, pages 1326–1331. IEEE, 2002.

[2] Maarten Keijzer. Improving Symbolic Regression with Interval Arithmetic and Linear Scaling. In *European Conference on Genetic Programming*, pages 70–82. Springer, 2003.

[3] John R Koza and Riccardo Poli. Genetic Programming. In *Search methodologies*, pages 127–164. Springer, 2005.

[9] If those differ by a constant or scalar, SRBench [11] treats the estimated equation as correct.

[4] Colin G Johnson. Genetic Programming Crossover: Does it Cross Over? In *European Conference on Genetic Programming*, pages 97–108. Springer, 2009.

[5] Nguyen Quang Uy, Nguyen Xuan Hoai, Michael O'Neill, Robert I McKay, and Edgar Galván-López. Semantically-based Crossover in Genetic Programming: Application to Real-valued Symbolic Regression. *Genetic Programming and Evolvable Machines*, 12(2):91–119, 2011.

[6] Patryk Orzechowski, William La Cava, and Jason H Moore. Where are we now? A large benchmark study of recent symbolic regression methods. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pages 1183–1190, 2018.

[7] Brenden K Petersen, Mikel Landajuela Larma, Terrell N Mundhenk, Claudio Prata Santiago, Soo Kyung Kim, and Joanne Taery Kim. Deep symbolic regression: Recovering mathematical expressions from data via risk-seeking policy gradients. In *International Conference on Learning Representations*, 2020.

[8] Mikel Landajuela, Brenden K Petersen, Sookyung Kim, Claudio P Santiago, Ruben Glatt, Nathan Mundhenk, Jacob F Pettit, and Daniel Faissol. Discovering symbolic policies with deep reinforcement learning. In *International Conference on Machine Learning*, pages 5979–5989. PMLR, 2021.

[9] Luca Biggio, Tommaso Bendinelli, Alexander Neitz, Aurelien Lucchi, and Giambattista Parascandolo. Neural Symbolic Regression that Scales. In *International Conference on Machine Learning*, pages 936–945. PMLR, 2021.

[10] Mojtaba Valipour, Bowen You, Maysum Panju, and Ali Ghodsi. SymbolicGPT: A Generative Transformer Model for Symbolic Regression. *arXiv preprint arXiv:2106.14131*, 2021.

[11] William La Cava, Patryk Orzechowski, Bogdan Burlacu, Fabricio Olivetti de Franca, Marco Virgolin, Ying Jin, Michael Kommenda, and Jason H Moore. Contemporary Symbolic Regression Methods and their Relative Performance. In *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 1)*, 2021.

[12] Pierre-Alexandre Kamienny, Stéphane d'Ascoli, Guillaume Lample, and François Charton. End-to-end symbolic regression with transformers. *arXiv preprint arXiv:2204.10532*, 2022.

[13] Tailin Wu and Max Tegmark. Toward an artificial intelligence physicist for unsupervised learning. *Physical Review E*, 100(3):033311, 2019.

[14] Silviu-Marian Udrescu and Max Tegmark. AI Feynman: A physics-inspired method for symbolic regression. *Science Advances*, 6(16):eaay2631, 2020.

[15] Silviu-Marian Udrescu, Andrew Tan, Jiahai Feng, Orisvaldo Neto, Tailin Wu, and Max Tegmark. AI Feynman 2.0: Pareto-optimal symbolic regression exploiting graph modularity. *Advances in Neural Information Processing Systems*, 33, 2020.

[16] Samuel Kim, Peter Y Lu, Srijon Mukherjee, Michael Gilbert, Li Jing, Vladimir Čeperić, and Marin Soljačić. Integration of Neural Network-Based Symbolic Regression in Deep Learning for Scientific Discovery. *IEEE Transactions on Neural Networks and Learning Systems*, 2020.

[17] Miles Cranmer, Alvaro Sanchez Gonzalez, Peter Battaglia, Rui Xu, Kyle Cranmer, David Spergel, and Shirley Ho. Discovering Symbolic Models from Deep Learning with Inductive Biases. *Advances in Neural Information Processing Systems*, 33:17429–17442, 2020.

[18] Ziming Liu and Max Tegmark. Machine Learning Conservation Laws from Trajectories. *Physical Review Letters*, 126:180604, May 2021.

[19] Z Liu, B Wang, Q Meng, W Chen, M Tegmark, and TY Liu. Machine-learning nonconservative dynamics for new-physics detection. *Physical Review. E*, 104(5-2):055302–055302, 2021.

[20] Zhanchao Huang, Chunjiang Li, Zhilong Huang, Yong Wang, and Hanqing Jiang. AI-Timoshenko: Automatedly Discovering Simplified Governing Equations for Applied Mechanics Problems From Simulated Data. *Journal of Applied Mechanics*, 88(10):101006, 2021.

[21] Ismail Alaoui Abdellaoui and Siamak Mehrkanoon. Symbolic regression for scientific discovery: an application to wind speed forecasting. In *2021 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 01–08. IEEE, 2021.

[22] Sheng Sun, Runhai Ouyang, Bochao Zhang, and Tong-Yi Zhang. Data-driven discovery of formulas by symbolic regression. *MRS Bulletin*, 44(7):559–564, 2019.

[23] Yiqun Wang, Nicholas Wagner, and James M Rondinelli. Symbolic regression in materials science. *MRS Communications*, 9(3):793–805, 2019.

[24] Baicheng Weng, Zhilong Song, Rilong Zhu, Qingyu Yan, Qingde Sun, Corey G Grice, Yanfa Yan, and Wan-Jian Yin. Simple descriptor derived from symbolic regression accelerating the discovery of new perovskite catalysts. *Nature Communications*, 11(1):1–8, 2020.

[25] Christian Loftis, Kunpeng Yuan, Yong Zhao, Ming Hu, and Jianjun Hu. Lattice Thermal Conductivity Prediction Using Symbolic Regression and Machine Learning. *The Journal of Physical Chemistry A*, 125(1):435–450, 2020.

[26] Rohit Batra, Le Song, and Rampi Ramprasad. Emerging materials intelligence ecosystems propelled by machine learning. *Nature Reviews Materials*, pages 1–24, 2020.

[27] Steven H Strogatz. *Nonlinear Dynamics and Chaos with Student Solutions Manual: With Applications to Physics, Biology, Chemistry, and Engineering*. CRC press, 2018.

[28] Herbert Goldstein, Charles Poole, and John Safko. *Classical Mechanics*. Addison Wesley, 2002.

[29] John David Jackson. *Classical Electrodynamics*. Wiley, 1999.

[30] Steven Weinberg. *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*. Wiley,, 1972.

[31] Matthew D Schwartz. *Quantum Field Theory and the Standard Model*. Cambridge University Press, 2014.

[32] Richard P Feynman, Robert B Leighton, and Matthew Sands. *The Feynman Lectures on Physics, Vol. I: The New Millennium Edition: Mainly Mechanics, Radiation, and Heat*, volume 1. Basic books, 1963.

[33] Richard P Feynman, Robert B Leighton, and Matthew Sands. *The Feynman Lectures on Physics, Vol. II: The New Millennium Edition: Mainly Electromagnetism and Matter*, volume 2. Basic books, 1963.

[34] Richard P Feynman, Robert B Leighton, and Matthew Sands. *The Feynman Lectures on Physics, Vol. II: The New Millennium Edition: Quantum Mechanics*, volume 3. Basic books, 1963.

[35] Michael Schmidt and Hod Lipson. Distilling Free-Form Natural Laws from Experimental Data. *Science*, 324(5923):81–85, 2009.

[36] William La Cava, Kourosh Danai, and Lee Spector. Inference of compact nonlinear dynamic models by epigenetic local search. *Engineering Applications of Artificial Intelligence*, 55: 292–306, 2016.

[37] National Astronomical Observatory of Japan. *Handbook of Scientific Tables*. World Scientific, 2022.

[38] Michael Schmidt and Hod Lipson. Age-Fitness Pareto Optimization. In *Genetic programming theory and practice VIII*, pages 129–146. Springer, 2011.

[39] Yoshitomo Matsubara, Naoya Chiba, Ryo Igarashi, Tatsunori Taniai, and Yoshitaka Ushiku. Rethinking Symbolic Regression Datasets and Benchmarks for Scientific Discovery. *arXiv preprint arXiv:2206.10540*, 2022.

[40] Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta, and Masanori Koyama. Optuna: A Next-generation Hyperparameter Optimization Framework. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 2623–2631, 2019.

A Our SRSD Datasets: Additional Information

This section provides additional information regarding our SRSD datasets. We created the datasets to discuss the performance of symbolic regression for scientific discovery (SRSD). We refer readers to Section 3 for details of the datasets. Tables S1 – S11 comprehensively summarize the differences between FSRD and our SRSD datasets. Note that the table of Easy set (part 1) is provided as Table 1 in Section 3.1. As described in Section 3.2.2, we categorized each of the 120 SRSD datasets into one of Easy, Medium, and Hard sets. We published the three sets of the SRSD datasets with MIT License at Hugging Face Dataset repositories. The dataset documentations are publicly available as Hugging Face Dataset cards.[1, 2, 3] These repositories are version-controlled with Git [10] so that users can track the log of the changes. We bear all responsibility in case of violation of rights.

Table S1: Easy set of our proposed datasets (part 2).

Eq. ID	Formula	Symbols		Properties		Distributions	
				Original	Ours	Original	Ours
I.30.5	$d = \frac{\lambda}{n \sin \theta}$	d	Interplanar distance	V, F	V, F, P	$\mathcal{U}(2, 5)$	N/A
		λ	Wavelength of X-ray	V, F	V, F, P	$\mathcal{U}(1, 2)$	$\mathcal{U}_{\log}(10^{-11}, 10^{-9})$
		n	The number of phase difference	V, F	V, I,P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{0}, 10^{2})$
		θ	Incidence/Reflection angle	V, F	V, F	N/A	$\mathcal{U}(-2\pi, 2\pi)$
I.43.16	$v = \mu q \frac{V}{d}$	v	Velocity	V, F	V, F	N/A	N/A
		μ	Ionic conductivity	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-6}, 10^{-4})$
		q	Electric charge of ions	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-11}, 10^{-9})$
		V	Voltage	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		d	Distance	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$
I.47.23	$c = \sqrt{\frac{\gamma P}{\rho}}$	c	Velocity of sound	V, F	V, F, P	N/A	N/A
		γ	Heat capacity ratio	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}(1, 2)$
		P	Atmospheric pressure	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}(0.5 \times 10^{-5}, 1.5 \times 10^{-5})$
		ρ	Density of air	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}(1, 2)$
II.2.42	$J = \kappa(T_2 - T_1)\frac{A}{d}$	J	Energy difference	V, F	V, F	N/A	N/A
		κ	Thermal conductivity	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		T_2	Temperature	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{1}, 10^{3})$
		T_1	Temperature	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{1}, 10^{3})$
		A	Area	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-4}, 10^{-2})$
		d	Length	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-2}, 10^{0})$
II.3.24	$h = \frac{W}{4\pi r^2}$	h	Heat flux	V, F	V, F	N/A	N/A
		W	Work	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{0}, 10^{2})$
		r	Distance	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-2}, 10^{0})$
II.4.23	$\phi = \frac{q}{4\pi\epsilon r}$	ϕ	Electric potential	V, F	V, F	N/A	N/A
		q	Electric charge	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$
		ϵ	Vacuum permittivity	V, F	C, F, P	$\mathcal{U}(1, 5)$	8.854×10^{-12}
		r	Distance	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-2}, 10^{0})$
II.8.31	$u = \frac{\epsilon E^2}{2}$	u	Energy	V, F	V, F	N/A	N/A
		ϵ	Vacuum permittivity	V, F	C, F, P	$\mathcal{U}(1, 5)$	8.854×10^{-12}
		E	Magnitude of electric field	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{1}, 10^{3})$
II.10.9	$E = \frac{\sigma_{\text{free}}}{\epsilon}\frac{1}{1+\chi}$	E	Electric field	V, F	V, F	N/A	N/A
		σ_{free}	Surface charge	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$
		ϵ	Vacuum permittivity	V, F	C, F, P	$\mathcal{U}(1, 5)$	8.854×10^{-12}
		χ	Electric susceptibility	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{0}, 10^{2})$
II.13.17	$B = \frac{1}{4\pi\epsilon c^2}\frac{2I}{r}$	B	The magnitude of the magnetic field	V, F	V, F	N/A	N/A
		ϵ	Vacuum permittivity	V, F	C, F, P	$\mathcal{U}(1, 5)$	8.854×10^{-12}
		c	Speed of light	V, F	C, F, P	$\mathcal{U}(1, 5)$	2.998×10^{8}
		I	Electric current	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$
		r	Radius	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$
II.15.4	$U = -\mu B \cos\theta$	U	Energy from magnetic field	V, F	V, F	N/A	N/A
		μ	Magnetic dipole moment	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-25}, 10^{-23})$
		B	Magnetic field strength	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$
		θ	Angle	V, F	V, F, NN	$\mathcal{U}(1, 5)$	$\mathcal{U}(0, 2\pi)$

[10] https://git-scm.com/

Table S2: Easy set of our proposed datasets (part 3).

Eq. ID	Formula	Symbols		Properties		Distributions	
				Original	Ours	Original	Ours
II.15.5	$U = -pE\cos\theta$	U	Energy	V, F	V, F	N/A	N/A
		p	Electric dipole moment	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-22}, 10^{-20})$
		E	Magnitude of electric field	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{1}, 10^{3})$
		θ	Angle	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}(0, 2\pi)$
II.27.16	$S = \epsilon c E^2$	S	Radiant intensity	V, F	V, F	N/A	N/A
		ϵ	Vacuum permittivity	V, F	C, F, P	$\mathcal{U}(1,5)$	8.854×10^{-12}
		c	Speed of light	V, F	C, F, P	$\mathcal{U}(1,5)$	2.998×10^{8}
		E	Magnitude of electric field	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
II.27.18	$u = \epsilon E^2$	u	Energy density	V, F	V, F, P	N/A	N/A
		ϵ	Vacuum permittivity	V, F	C, F, P	$\mathcal{U}(1,5)$	8.854×10^{-12}
		E	Magnitude of electric field	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
II.34.11	$\omega = g\frac{qB}{2m}$	ω	Angular frequency	V, F	V, F	N/A	N/A
		g	g-factor	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}(-1, 1)$
		q	Electric charge	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-11}, 10^{-9})$
		B	Magnetic field strength	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-9}, 10^{-7})$
		m	Mass	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-30}, 10^{-28})$
II.34.29b	$U = 2\pi g\mu B\frac{J_z}{h}$	U	Energy	V, F	V, F, P	N/A	N/A
		g	g-factor	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}(-1, 1)$
		μ	Bohr magneton	V, F	C, F, P	$\mathcal{U}(1,5)$	$9.2740100783 \times 10^{-24}$
		B	Magnetic field strength	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$
		J_z	Element of angular momentum	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-26}, 10^{-22})$
		h	Planck constant	V, F	C, F, P	$\mathcal{U}(1,5)$	6.626×10^{-34}
II.38.3	$F = YA\frac{\Delta l}{l}$	F	Force	V, F	V, F	N/A	N/A
		Y	Young's modulus	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		A	Area	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-4}, 10^{-2})$
		δl	Displacement	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$
		l	Length	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-2}, 10^{0})$
II.38.14	$\mu = \frac{Y}{2(1+\sigma)}$	μ	Rigidity modulus	V, F	V, F, P	N/A	N/A
		Y	Young's modulus	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		σ	Poisson coefficient	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-2}, 10^{0})$
III.7.38	$\omega = \frac{4\pi\mu B}{h}$	ω	Precession frequency	V, F	V, F	N/A	N/A
		μ	Magnetic moment	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-11}, 10^{-9})$
		B	Magnetic flux density	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$
		h	Planck constant	V, F	C, F, P	$\mathcal{U}(1,5)$	6.626×10^{-34}
III.12.43	$J = \frac{mh}{2\pi}$	J	Variable	V, F	V, F	N/A	N/A
		m	Spin state	V, F	V, I,NN	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{0}, 10^{2})$
		h	Planck constant	V, F	C, F, P	$\mathcal{U}(1,5)$	6.626×10^{-34}
III.15.27	$k = \frac{2\pi}{Nb}s$	k	Wavenumber	V, F	V, F	N/A	N/A
		s	Parameter of state	V, F	V, I	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{0}, 10^{2})$
		N	Number of atoms	V, F	V, I,P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{0}, 10^{2})$
		b	Lattice constant	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-10}, 10^{-8})$

Table S3: Medium set of our proposed datasets (part 1). C: Constant, V: Variable, F: Float, I: Integer, P: Positive, N: Negative, NN: Non-Negative, I$\star$: Integer treated as float due to the capacity of 32-bit integer.

Eq. ID	Formula	Symbols		Properties		Distributions	
				Original	Ours	Original	Ours
I.8.14	$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$	d	Distance	V, F	V, F, NN	N/A	N/A
		x_2	Position	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		x_1	Position	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		y_2	Position	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		y_1	Position	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
I.10.7	$m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}$	m	Mass	V, F	V, F, P	N/A	N/A
		m_0	Invariant mass	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		v	Velocity	V, F	V, F, P	$\mathcal{U}(1, 2)$	$\mathcal{U}_{\log}(10^{5}, 10^{8})$
		c	Speed of light	V, F	C, F, P	$\mathcal{U}(3, 10)$	2.998×10^{8}
I.11.19	$A = x_1 y_1 + x_2 y_2 + x_3 y_3$	A	Inner product	V, F	V, F	N/A	N/A
		x_1	Element of a vector	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		y_1	Element of a vector	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		x_2	Element of a vector	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		y_2	Element of a vector	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		x_3	Element of a vector	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		y_3	Element of a vector	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
I.12.2	$F = \frac{q_1 q_2}{4 \pi \epsilon r^2}$	F	Electrostatic force	V, F	V, F	N/A	N/A
		q_1	Electric charge	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$
		q_2	Electric charge	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$
		r	Distance	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-2}, 10^{0})$
		ϵ	Vacuum permittivity	V, F	C, F, P	$\mathcal{U}(1, 5)$	8.854×10^{-12}
I.12.11	$F = q\left(E + Bv \sin(\theta)\right)$	F	Force	V, F	V, F	N/A	N/A
		q	Electric charge	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		E	Electric field	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		B	Magnetic field strength	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		v	Velocity	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		θ	Angle	V, F	V, F, NN	$\mathcal{U}(1, 5)$	$\mathcal{U}(0, 2\pi)$
I.13.4	$K = \frac{1}{2} m (v^2 + u^2 + w^2)$	K	Kinetic energy	V, F	V, F, P	N/A	N/A
		m	Mass	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-2}, 10^{0})$
		v	Element of velocity	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		u	Element of velocity	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		w	Element of velocity	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
I.13.12	$U = G m_1 m_2 \left(\frac{1}{r_2} - \frac{1}{r_1}\right)$	U	Potential energy	V, F	V, F, P	N/A	N/A
		G	Gravitational constant	V, F	C, F, P	$\mathcal{U}(1, 5)$	6.674×10^{-11}
		m_1	Mass (The Earth)	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-2}, 10^{0})$
		m_2	Mass	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-2}, 10^{0})$
		r_2	Distance	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-2}, 10^{0})$
		r_1	Distance	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-2}, 10^{0})$
I.15.10	$p = \frac{m_0 v}{\sqrt{1-v^2/c^2}}$	p	Relativistic mass	V, F	V, F, P	N/A	N/A
		m_0	Rest Mass	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-2}, 10^{0})$
		v	Velocity	V, F	V, F	$\mathcal{U}(1, 2)$	$\mathcal{U}_{\log}(10^{5}, 10^{7})$
		c	Speed of light	V, F	C, F, P	$\mathcal{U}(3, 10)$	2.998×10^{8}
I.16.6	$v_1 = \frac{u+v}{1+uv/c^2}$	v_1	Velocity	V, F	V, F	N/A	N/A
		u	Velocity	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{6}, 10^{8})$
		v	Velocity	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{6}, 10^{8})$
		c	Speed of light	V, F	C, F, P	$\mathcal{U}(1, 5)$	2.998×10^{8}
I.18.4	$r = \frac{m_1 r_1 + m_2 r_2}{m_1 + m_2}$	r	Center of gravity	V, F	V, F	N/A	N/A
		m_1	Mass	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		r_1	Position	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		m_2	Mass	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		r_2	Position	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$

Table S4: Medium set of our proposed datasets (part 2).

Eq. ID	Formula	Symbols		Properties		Distributions	
				Original	Ours	Original	Ours
I.24.6	$E = \frac{1}{4}m(\omega^2 + \omega_0^2)x^2$	E	Energy	V, F	V, F, P	N/A	N/A
		m	Mass	V, F	V, F, P	$\mathcal{U}(1, 3)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		ω	Angular velocity	V, F	V, F	$\mathcal{U}(1, 3)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		ω_0	Angular velocity	V, F	V, F	$\mathcal{U}(1, 3)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		x	Position	V, F	V, F	$\mathcal{U}(1, 3)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
I.29.4	$k = \frac{\omega}{c}$	k	Wavenumber	V, F	V, F, P	N/A	N/A
		ω	Frequency of electromagnetic waves	V, F	V, F, P	$\mathcal{U}(1, 10)$	$\mathcal{U}_{\log}(10^{9}, 10^{11})$
		c	Speed of light	V, F	C, F, P	$\mathcal{U}(1, 10)$	2.998×10^{8}
I.32.5	$P = \frac{q^2 a^2}{6\pi\epsilon c^3}$	P	Radiant energy	V, F	V, F, P	N/A	N/A
		q	Electric charge	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$
		a	Magnitude of direction vector	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{5}, 10^{7})$
		ϵ	Vacuum permittivity	V, F	C, F, P	$\mathcal{U}(1, 5)$	8.854×10^{-12}
		c	Speed of light	V, F	C, F, P	$\mathcal{U}(1, 5)$	2.998×10^{8}
I.34.8	$\omega = \frac{qvB}{p}$	ω	Angular velocity	V, F	V, F	N/A	N/A
		q	Electric charge	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-11}, 10^{-9})$
		v	Velocity	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{5}, 10^{7})$
		B	Magnetic field	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{1}, 10^{3})$
		p	Angular momentum	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{9}, 10^{11})$
I.34.10	$\omega = \frac{\omega_0}{1 - v/c}$	ω	Frequency of electromagnetic waves	V, F	V, F, P	N/A	N/A
		ω_0	Frequency of electromagnetic waves	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{9}, 10^{11})$
		v	Velocity	V, F	V, F	$\mathcal{U}(1, 2)$	$\mathcal{U}_{\log}(10^{5}, 10^{7})$
		c	Speed of light	V, F	C, F, P	$\mathcal{U}(3, 10)$	2.998×10^{8}
I.34.27	$W = \frac{h}{2\pi}\omega$	W	Energy	V, F	V, F, P	N/A	N/A
		h	Planck constant	V, F	C, F, P	$\mathcal{U}(1, 5)$	6.626×10^{-34}
		ω	Frequency of electromagnetic waves	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{9}, 10^{11})$
I.38.12	$r = 4\pi\epsilon \frac{(h/(2\pi))^2}{mq^2}$	r	Bohr radius	V, F	V, F, P	N/A	N/A
		ϵ	Vacuum permittivity	V, F	C, F, P	$\mathcal{U}(1, 5)$	8.854×10^{-12}
		h	Planck constant	V, F	C, F, P	$\mathcal{U}(1, 5)$	6.626×10^{-34}
		m	Mass	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-28}, 10^{-26})$
		q	Electric charge	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-11}, 10^{-9})$
I.39.10	$U = \frac{3}{2}PV$	U	Internal energy	V, F	V, F, P	N/A	N/A
		P	Pressure	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{4}, 10^{6})$
		V	Volume	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-5}, 10^{-3})$
I.39.11	$U = \frac{PV}{\gamma - 1}$	U	Energy	V, F	V, F	N/A	N/A
		γ	Heat capacity ratio	V, F	V, F, P	$\mathcal{U}(2, 5)$	$\mathcal{U}(1, 2)$
		P	Pressure	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{4}, 10^{6})$
		V	Volume	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-5}, 10^{-3})$
I.43.31	$D = \mu kT$	D	Diffusion coefficient	V, F	V, F, P	N/A	N/A
		μ	Viscosity	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{13}, 10^{15})$
		k	Boltzmann constant	V, F	C, F, P	$\mathcal{U}(1, 5)$	1.381×10^{-23}
		T	Temperature	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{1}, 10^{3})$

Table S5: Medium set of our proposed datasets (part 3).

Eq. ID	Formula	Symbols		Properties		Distributions	
				Original	Ours	Original	Ours
I.43.43	$\kappa = \frac{1}{\gamma-1}\frac{kv}{\sigma_c}$	κ	Thermal conductivity	V, F	V, F, P	N/A	N/A
		γ	Heat capacity ratio	V, F	V, F, P	$\mathcal{U}(2,5)$	$\mathcal{U}(1,2)$
		k	Boltzmann constant	V, F	C, F, P	$\mathcal{U}(1,5)$	1.381×10^{-23}
		v	Velocity	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{2}, 10^{4})$
		σ_c	Molecular collision cross section	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-21}, 10^{-19})$
I.48.2	$E = \frac{mc^2}{\sqrt{1-v^2/c^2}}$	E	Energy	V, F	V, F, P	N/A	N/A
		m	Mass	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-29}, 10^{-27})$
		c	Speed of light	V, F	C, F, P	$\mathcal{U}(3,10)$	2.998×10^{8}
		v	Velocity	V, F	V, F, P	$\mathcal{U}(1,2)$	$\mathcal{U}_{\log}(10^{6}, 10^{8})$
II.6.11	$\phi = \frac{1}{4\pi\epsilon}\frac{p\cos\theta}{r^2}$	ϕ	Electric potential	V, F	V, F	N/A	N/A
		ϵ	Vacuum permittivity	V, F	C, F, P	$\mathcal{U}(1,3)$	8.854×10^{-12}
		p	Electric dipole moment	V, F	V, F	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-22}, 10^{-20})$
		θ	Angle	V, F	V, F, NN	$\mathcal{U}(1,3)$	$\mathcal{U}(0, 2\pi)$
		r	Distance	V, F	V, F, P	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-10}, 10^{-8})$
II.8.7	$U = \frac{3}{5}\frac{Q^2}{4\pi\epsilon a}$	U	Energy	V, F	V, F	N/A	N/A
		Q	Electric charge	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-11}, 10^{-9})$
		ϵ	Vacuum permittivity	V, F	C, F, P	$\mathcal{U}(1,5)$	8.854×10^{-12}
		a	Radius	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-12}, 10^{-10})$
II.11.3	$x = \frac{qE}{m(\omega_0^2-\omega^2)}$	x	Position	V, F	V, F	N/A	N/A
		q	Electric charge	V, F	V, F	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-11}, 10^{-9})$
		E	Magnitude of electric field	V, F	V, F, P	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-9}, 10^{-7})$
		m	Mass	V, F	V, F, P	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-28}, 10^{-26})$
		ω_0	Angular velocity	V, F	V, F	$\mathcal{U}(3,5)$	$\mathcal{U}_{\log}(10^{9}, 10^{11})$
		ω	Angular velocity	V, F	V, F	$\mathcal{U}(1,2)$	$\mathcal{U}_{\log}(10^{9}, 10^{11})$
II.21.32	$\phi = \frac{q}{4\pi\epsilon r(1-v/c)}$	ϕ	Electric potential	V, F	V, F	N/A	N/A
		q	Electric charge	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$
		ϵ	Vacuum permittivity	V, F	C, F, P	$\mathcal{U}(1,5)$	8.854×10^{-12}
		r	Distance	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{0}, 10^{2})$
		v	Velocity	V, F	V, F, P	$\mathcal{U}(1,2)$	$\mathcal{U}_{\log}(10^{5}, 10^{7})$
		c	Speed of light	V, F	C, F, P	$\mathcal{U}(3,10)$	2.998×10^{8}
II.34.2	$\mu = \frac{qvr}{2}$	μ	Magnetic moment	V, F	V, F	N/A	N/A
		q	Electric charge	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-11}, 10^{-9})$
		v	Velocity	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{5}, 10^{7})$
		r	Radius	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-11}, 10^{-9})$
II.34.2a	$I = \frac{qv}{2\pi r}$	I	Electric Current	V, F	V, F	N/A	N/A
		q	Electric charge	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-11}, 10^{-9})$
		v	Velocity	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{5}, 10^{7})$
		r	Radius	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-11}, 10^{-9})$
II.34.29a	$\mu = \frac{qh}{4\pi m}$	μ	Bohr magneton	V, F	V, F	N/A	N/A
		q	Electric charge	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-11}, 10^{-9})$
		h	Planck constant	V, F	C, F, P	$\mathcal{U}(1,5)$	6.626×10^{-34}
		m	Mass	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-30}, 10^{-28})$
II.37.1	$E = \mu(1+\chi)B$	E	Energy of magnetic field	V, F	V, F	N/A	N/A
		μ	Magnetic moment	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-25}, 10^{-23})$
		χ	Volume magnetic susceptibility	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{4}, 10^{6})$
		B	Magnetic field strength	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$

Table S6: Medium set of our proposed datasets (part 4).

Eq. ID	Formula	Symbols		Properties		Distributions	
				Original	Ours	Original	Ours
III.4.32	$n = \frac{1}{\exp(h\omega/2\pi kT)-1}$	n	Average number of photons	V, F	V, F, P	N/A	N/A
		h	Planck constant	V, F	C, F, P	$\mathcal{U}(1,5)$	6.626×10^{-34}
		ω	Frequency	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{8}, 10^{10})$
		k	Boltzmann constant	V, F	C, F, P	$\mathcal{U}(1,5)$	1.381×10^{-23}
		T	Temperature	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{1}, 10^{3})$
III.8.54	$\lvert C\rvert^2 = \sin^2\left(\frac{2\pi At}{h}\right)$	$\lvert C\rvert^2$	Probability	V, F	V, F, NN	N/A	N/A
		A	Energy	V, F	V, F	$\mathcal{U}(1,2)$	$\mathcal{U}_{\log}(10^{-18}, 10^{-16})$
		t	Time	V, F	V, F, NN	$\mathcal{U}(1,2)$	$\mathcal{U}_{\log}(10^{-18}, 10^{-16})$
		h	Planck constant	V, F	C, F, P	$\mathcal{U}(1,4)$	6.626×10^{-34}
III.13.18	$v = \frac{4\pi Ab^2}{h}k$	v	Speed of the waves	V, F	V, F	N/A	N/A
		A	Energy	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-18}, 10^{-16})$
		b	Lattice constant	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-10}, 10^{-8})$
		k	Wavenumber	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		h	Planck constant	V, F	C, F, P	$\mathcal{U}(1,5)$	6.626×10^{-34}
III.14.14	$I = I_0\left(\exp\left(q\Delta V/\kappa T\right)-1\right)$	I	Electric Current	V, F	V, F	N/A	N/A
		I_0	Electric current	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$
		q	Electric charge	V, F	V, F, P	$\mathcal{U}(1,2)$	$\mathcal{U}_{\log}(10^{-22}, 10^{-20})$
		ΔV	Voltage	V, F	V, F	$\mathcal{U}(1,2)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		κ	Boltzmann constant	V, F	C, F, P	$\mathcal{U}(1,2)$	1.381×10^{-23}
		T	Temperature	V, F	V, F, P	$\mathcal{U}(1,2)$	$\mathcal{U}_{\log}(10^{1}, 10^{3})$
III.15.12	$E = 2A\left(1-\cos\left(kd\right)\right)$	E	Energy	V, F	V, F, P	N/A	N/A
		A	Amplitude	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-18}, 10^{-16})$
		k	Propagation coefficient	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		d	Lattice constant	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-10}, 10^{-8})$
III.15.14	$m = \frac{h^2}{8\pi^2 Ab^2}$	m	Effective mass	V, F	V, F, P	N/A	N/A
		h	Planck constant	V, F	C, F, P	$\mathcal{U}(1,5)$	6.626×10^{-34}
		A	Amplitude	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-18}, 10^{-16})$
		b	Lattice constant	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-10}, 10^{-8})$
III.17.37	$f = \beta(1+\alpha\cos\theta)$	f	Distribution	V, F	V, F	N/A	N/A
		β	Variable	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-18}, 10^{-16})$
		α	Variable	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-18}, 10^{-16})$
		θ	Angle	V, F	V, F, NN	$\mathcal{U}(1,5)$	$\mathcal{U}(0, 2\pi)$
III.19.51	$E = -\frac{mq^4}{2(4\pi\epsilon)^2(h/(2\pi))^2 n^2}$	E	Energy	V, F	V, F, P	N/A	N/A
		m	Mass	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-30}, 10^{-28})$
		q	Electric charge	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-11}, 10^{-9})$
		ϵ	Vacuum permittivity	V, F	C, F, P	$\mathcal{U}(1,5)$	8.854×10^{-12}
		h	Planck constant	V, F	C, F, P	$\mathcal{U}(1,5)$	6.626×10^{-34}
		n	Number of protons	V, F	V, I,P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{0}, 10^{2})$
B8	$U = \frac{E}{1+\frac{E}{mc^2}(1-\cos\theta)}$	U	Variable	V, F	V, F, P	N/A	N/A
		E	Electromagnetic energy	V, F	V, F, P	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-24}, 10^{-22})$
		m	Electron mass	V, F	C, F, P	$\mathcal{U}(1,3)$	9.109×10^{-31}
		c	Speed of light	V, F	C, F, P	$\mathcal{U}(1,3)$	2.998×10^{8}
		θ	Incidence angle	V, F	V, F	$\mathcal{U}(1,3)$	$\mathcal{U}(-\pi, \pi)$
B18	$\rho = \frac{3}{8\pi G}\left(\frac{c^2 k_\mathrm{f}}{a_\mathrm{f}^2}+H^2\right)$	ρ	Variable	V, F	V, F	N/A	N/A
		G	Gravitational constant	V, F	C, F, P	$\mathcal{U}(1,5)$	6.674×10^{-11}
		c	Speed of light	V, F	C, F, P	$\mathcal{U}(1,5)$	2.998×10^{8}
		k_f	Variable	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{1}, 10^{3})$
		a_f	Distance	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{8}, 10^{10})$
		H	Variable	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{0}, 10^{2})$

Table S7: Hard set of our proposed datasets (part 1).

Eq. ID	Formula	Symbols		Properties Original	Properties Ours	Distributions Original	Distributions Ours
I.6.20	$f = \exp\left(-\frac{\theta^2}{2\sigma^2}\right)/\sqrt{2\pi\sigma^2}$	f	Probability density function	V, F	V, F	N/A	N/A
		θ	Position	V, F	V, F	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-1},10^{1})$
		σ	Standard deviation	V, F	V, F, P	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-1},10^{1})$
I.6.20a	$f = \exp\left(-\frac{\theta^2}{2}\right)/\sqrt{2\pi}$	f	Probability density function	V, F	V, F	N/A	N/A
		θ	Position	V, F	V, F	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-1},10^{1})$
I.6.20b	$f = \exp\left(-\frac{(\theta-\theta_1)^2}{2\sigma^2}\right)/\sqrt{2\pi\sigma}$	f	Probability density function	V, F	V, F	N/A	N/A
		θ	Position	V, F	V, F	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-1},10^{1})$
		θ_1	Position	V, F	V, F	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-1},10^{1})$
		σ	Standard deviation	V, F	V, F, P	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-1},10^{1})$
I.9.18	$F = \frac{Gm_1m_2}{(x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2}$	F	Force of gravity	V, F	V, F	N/A	N/A
		G	Gravitational constant	V, F	C, F, P	$\mathcal{U}(1,2)$	6.674×10^{-11}
		m_1	Mass	V, F	V, F, P	$\mathcal{U}(1,2)$	$\mathcal{U}_{\log}(10^{0},10^{3})$
		m_2	Mass	V, F	V, F, P	$\mathcal{U}(1,2)$	$\mathcal{U}_{\log}(10^{0},10^{3})$
		x_2	Position	V, F	V, F	$\mathcal{U}(1,2)$	$\mathcal{U}_{\log}(10^{0},10^{1})$
		x_1	Position	V, F	V, F	$\mathcal{U}(3,4)$	$\mathcal{U}_{\log}(10^{0},10^{1})$
		y_2	Position	V, F	V, F	$\mathcal{U}(1,2)$	$\mathcal{U}_{\log}(10^{0},10^{1})$
		y_1	Position	V, F	V, F	$\mathcal{U}(3,4)$	$\mathcal{U}_{\log}(10^{0},10^{1})$
		z_2	Position	V, F	V, F	$\mathcal{U}(1,2)$	$\mathcal{U}_{\log}(10^{0},10^{1})$
		z_1	Position	V, F	V, F	$\mathcal{U}(3,4)$	$\mathcal{U}_{\log}(10^{0},10^{1})$
I.15.3t	$t_1 = \frac{t-ux/c^2}{\sqrt{1-u^2/c^2}}$	t_1	Time	V, F	V, F	N/A	N/A
		t	Time	V, F	V, F, NN	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-6},10^{-4})$
		u	Velocity	V, F	V, F	$\mathcal{U}(1,2)$	$\mathcal{U}_{\log}(10^{5},10^{7})$
		x	Position	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{0},10^{2})$
		c	Speed of light	V, F	C, F, P	$\mathcal{U}(3,10)$	2.998×10^{8}
I.15.3x	$x_1 = \frac{x-ut}{\sqrt{1-u^2/c^2}}$	x_1	Position	V, F	V, F	N/A	N/A
		x	Position	V, F	V, F	$\mathcal{U}(5,10)$	$\mathcal{U}_{\log}(10^{0},10^{2})$
		u	Velocity	V, F	V, F	$\mathcal{U}(1,2)$	$\mathcal{U}_{\log}(10^{6},10^{8})$
		t	Time	V, F	V, F	$\mathcal{U}(1,2)$	$\mathcal{U}_{\log}(10^{-6},10^{-4})$
		c	Speed of light	V, F	C, F, P	$\mathcal{U}(3,20)$	2.998×10^{8}
I.29.16	$x = \sqrt{x_1^2+x_2^2+2x_1x_2\cos(\theta_1-\theta_2)}$	x	Wavelength	V, F	V, F, P	N/A	N/A
		x_1	Wavelength	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-1},10^{1})$
		x_2	Wavelength	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-1},10^{1})$
		θ_1	Angle	V, F	V, F, NN	$\mathcal{U}(1,5)$	$\mathcal{U}(0,2\pi)$
		θ_2	Angle	V, F	V, F, NN	$\mathcal{U}(1,5)$	$\mathcal{U}(0,2\pi)$
I.30.3	$I = I_0\frac{\sin^2(n\theta/2)}{\sin^2(\theta/2)}$	I	Amplitude of combined wave	V, F	V, F	N/A	N/A
		I_0	Amplitude of wave	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-3},10^{-1})$
		n	The number of waves	V, F	V, I,P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{1},10^{3})$
		θ	Phase difference	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}(-2\pi,2\pi)$
I.32.17	$P = \left(\frac{1}{2}\epsilon cE^2\right)\left(\frac{8\pi r^2}{3}\right)\left(\frac{\omega^4}{(\omega^2-\omega_0^2)^2}\right)$	P	Energy	V, F	V, F, P	N/A	N/A
		ϵ	Vacuum permittivity	V, F	C, F, P	$\mathcal{U}(1,2)$	8.854×10^{-12}
		c	Speed of light	V, F	C, F, P	$\mathcal{U}(1,2)$	2.998×10^{8}
		E	Magnitude of electric field	V, F	V, F, P	$\mathcal{U}(1,2)$	$\mathcal{U}_{\log}(10^{1},10^{3})$
		r	Radius	V, F	V, F, P	$\mathcal{U}(1,2)$	$\mathcal{U}_{\log}(10^{-2},10^{0})$
		ω	Frequency of electromagnetic waves	V, F	V, F	$\mathcal{U}(1,2)$	$\mathcal{U}_{\log}(10^{9},10^{11})$
		ω_0	Frequency of electromagnetic waves	V, F	V, F	$\mathcal{U}(3,5)$	$\mathcal{U}_{\log}(10^{9},10^{11})$
I.34.14	$\omega = \frac{1+v/c}{\sqrt{1-v^2/c^2}}\omega_0$	ω	Frequency of electromagnetic waves	V, F	V, F	N/A	N/A
		v	Velocity	V, F	V, F	$\mathcal{U}(1,2)$	$\mathcal{U}_{\log}(10^{6},10^{8})$
		c	Speed of light	V, F	C, F, P	$\mathcal{U}(3,10)$	2.998×10^{8}
		ω_0	Frequency of electromagnetic waves	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{9},10^{11})$

Table S8: Hard set of our proposed datasets (part 2).

Eq. ID	Formula	Symbols		Properties		Distributions	
				Original	Ours	Original	Ours
I.37.4	$I_{12} = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos\delta$	I_{12}	Amplitude of wave	V, F	V, F, P	N/A	N/A
		I_1	Amplitude of wave	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{-3})$
		I_2	Amplitude of wave	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{-3})$
		δ	Phase difference	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}(0,\pi)$
I.39.22	$P = \frac{nkT}{V}$	P	Pressure	V, F	V, F, P	N/A	N/A
		n	Number of molecules	V, F	V, I⋆, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{23}, 10^{25})$
		k	Boltzmann constant	V, F	C, F, P	$\mathcal{U}(1,5)$	1.381×10^{-23}
		T	Temperature	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{1}, 10^{3})$
		V	Volume	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-5}, 10^{-3})$
I.40.1	$n = n_0 \exp(-mgx/kT)$	n	Molecular density	V, F	V, F, P	N/A	N/A
		n_0	Molecular density	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{25}, 10^{27})$
		m	Mass	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-24}, 10^{-22})$
		g	Gravitational acceleration	V, F	C, F, P	$\mathcal{U}(1,5)$	9.807×10^{0}
		x	Height	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-2}, 10^{0})$
		k	Boltzmann constant	V, F	C, F, P	$\mathcal{U}(1,5)$	1.381×10^{-23}
		T	Temperature	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{1}, 10^{3})$
I.41.16	$L_{\text{rad}} = \frac{h}{2\pi} \frac{\omega^3}{\pi^2 c^2 (\exp(h\omega/2\pi kT) - 1)}$	L_{rad}	Radiation per frequency	V, F	V, F, P	N/A	N/A
		h	Planck constant	V, F	C, F, P	$\mathcal{U}(1,5)$	6.626×10^{-34}
		ω	Frequency of electromagnetic wave	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		c	Speed of light	V, F	C, F, P	$\mathcal{U}(1,5)$	2.998×10^{8}
		k	Boltzmann constant	V, F	C, F, P	$\mathcal{U}(1,5)$	1.381×10^{-23}
		T	Temperature	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{1}, 10^{3})$
I.44.4	$Q = nkT \ln(\frac{V_2}{V_1})$	Q	Energy	V, F	V, F	N/A	N/A
		n	Number of molecules	V, F	V, I⋆, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{23}, 10^{25})$
		k	Boltzmann constant	V, F	C, F, P	$\mathcal{U}(1,5)$	1.381×10^{-23}
		T	Temperature	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{1}, 10^{3})$
		V_2	Volume	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-5}, 10^{-3})$
		V_1	Volume	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-5}, 10^{-3})$
I.50.26	$x = K\left(\cos\omega t + \epsilon\cos^2\omega t\right)$	x	Amplitude	V, F	V, F	N/A	N/A
		K	Amplitude	V, F	V, F, P	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		ω	Angular velocity	V, F	V, F	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{1}, 10^{3})$
		t	Time	V, F	V, F, NN	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$
		ϵ	Variable	V, F	V, F	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$
II.6.15a	$E = \frac{p}{4\pi\epsilon}\frac{3z}{r^5}\sqrt{x^2+y^2}$	E	Electric field	V, F	V, F	N/A	N/A
		p	Electric dipole moment	V, F	V, F	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-22}, 10^{-20})$
		ϵ	Vacuum permittivity	V, F	C, F, P	$\mathcal{U}(1,3)$	8.854×10^{-12}
		z	Position	V, F	V, F	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-10}, 10^{-8})$
		r	Distance	V, F	V, F, P	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-10}, 10^{-8})$
		x	Position	V, F	V, F	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-10}, 10^{-8})$
		y	Position	V, F	V, F	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-10}, 10^{-8})$
II.6.15b	$E = \frac{p}{4\pi\epsilon}\frac{3\cos\theta\sin\theta}{r^3}$	E	Electric field	V, F	V, F	N/A	N/A
		p	Electric dipole moment	V, F	V, F	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-22}, 10^{-20})$
		ϵ	Vacuum permittivity	V, F	C, F, P	$\mathcal{U}(1,3)$	8.854×10^{-12}
		θ	Angle	V, F	V, F	$\mathcal{U}(1,3)$	$\mathcal{U}(0,\pi)$
		r	Distance	V, F	V, F, P	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-10}, 10^{-8})$
II.11.17	$n = n_0\left(1 + \frac{p_0 E\cos\theta}{kT}\right)$	n	Number of polar molecules per angle per unit volume	V, F	V, F	N/A	N/A
		n_0	Number of molecules per unit volume	V, F	V, F, P	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{27}, 10^{29})$
		p_0	Electric dipole moment	V, F	V, F	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-22}, 10^{-20})$
		E	Magnitude of electric field	V, F	V, F	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{1}, 10^{3})$
		θ	Angle	V, F	V, F, NN	$\mathcal{U}(1,3)$	$\mathcal{U}(0,2\pi)$
		k	Boltzmann constant	V, F	C, F, P	$\mathcal{U}(1,3)$	1.381×10^{-23}
		T	Temperature	V, F	V, F, P	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{1}, 10^{3})$
II.11.20	$P = \frac{n_0 p_0^2 E}{3kT}$	P	Polarizability	V, F	V, F	N/A	N/A
		n_0	Number of atom	V, F	V, I⋆, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{23}, 10^{25})$
		p_0	Electric dipole moment	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-22}, 10^{-20})$
		E	Magnitude of electric field	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{1}, 10^{3})$
		k	Boltzmann constant	V, F	C, F, P	$\mathcal{U}(1,5)$	1.381×10^{-23}
		T	Temperature	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{1}, 10^{3})$

Table S9: Hard set of our proposed datasets (part 3).

Eq. ID	Formula	Symbols		Properties		Distributions	
				Original	Ours	Original	Ours
II.11.27	$P = \frac{N\alpha}{1-(n\alpha/3)}\epsilon E$	P	Polarizability	V, F	V, F	N/A	N/A
		N	Number of atom	V, F	V, I⋆, P	$\mathcal{U}(0,1)$	$\mathcal{U}_{\log}(10^{23},10^{25})$
		α	Molecular polarizability	V, F	V, F, P	$\mathcal{U}(0,1)$	$\mathcal{U}_{\log}(10^{-33},10^{-31})$
		ϵ	Vacuum permittivity	V, F	C, F, P	$\mathcal{U}(1,2)$	8.854×10^{-12}
		E	Magnitude of electric field	V, F	V, F, P	$\mathcal{U}(1,2)$	$\mathcal{U}_{\log}(10^{1},10^{3})$
II.11.28	$\kappa = 1 + \frac{N\alpha}{1-(N\alpha/3)}$	κ	Electric dipole moment per unit volume	V, F	V, F	N/A	N/A
		N	Number of electric dipoles	V, F	V, I⋆, P	$\mathcal{U}(0,1)$	$\mathcal{U}_{\log}(10^{23},10^{25})$
		α	Molecular polarizability	V, F	V, F, P	$\mathcal{U}(0,1)$	$\mathcal{U}_{\log}(10^{-33},10^{-31})$
II.13.23	$\rho = \frac{\rho_0}{\sqrt{1-v^2/c^2}}$	ρ	Electric charge density	V, F	V, F, P	N/A	N/A
		ρ_0	Electric charge density	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{27},10^{29})$
		v	Velocity	V, F	V, F, P	$\mathcal{U}(1,2)$	$\mathcal{U}_{\log}(10^{6},10^{8})$
		c	Speed of light	V, F	C, F, P	$\mathcal{U}(3,10)$	2.998×10^{8}
II.13.34	$j = \frac{\rho_0 v}{\sqrt{1-v^2/c^2}}$	j	Electric current	V, F	V, F	N/A	N/A
		ρ_0	Electric charge density	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{27},10^{29})$
		v	Velocity	V, F	V, F, P	$\mathcal{U}(1,2)$	$\mathcal{U}_{\log}(10^{6},10^{8})$
		c	Speed of light	V, F	C, F, P	$\mathcal{U}(3,10)$	2.998×10^{8}
II.24.17	$k = \sqrt{\omega^2/c^2 - \pi^2/a^2}$	k	Wavenumber	V, F	V, F, P	N/A	N/A
		ω	Angular velocity	V, F	V, F	$\mathcal{U}(4,6)$	$\mathcal{U}_{\log}(10^{9},10^{11})$
		c	Speed of light	V, F	C, F, P	$\mathcal{U}(1,2)$	2.998×10^{8}
		a	Length	V, F	V, F, P	$\mathcal{U}(2,4)$	$\mathcal{U}_{\log}(10^{-3},10^{-1})$
II.35.18	$a = \frac{N}{\exp(\mu B/kT)+\exp(-\mu B/kT)}$	a	Number of atoms with the equivalent magnetic moment	V, F	V, I⋆, P	N/A	N/A
		N	Number of atoms per unit volume	V, F	V, I⋆, P	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{23},10^{25})$
		μ	Magnetic moment	V, F	V, F, P	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-25},10^{-23})$
		B	Magnetic flux density	V, F	V, F, P	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-3},10^{-1})$
		k	Boltzmann constant	V, F	C, F, P	$\mathcal{U}(1,3)$	1.381×10^{-23}
		T	Temperature	V, F	V, F, P	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{1},10^{3})$
II.35.21	$M = N\mu\tanh\left(\frac{\mu B}{kT}\right)$	M	Number of magnetized atoms	V, F	V, I⋆, P	N/A	N/A
		N	Number of atom	V, F	V, I⋆, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{23},10^{25})$
		μ	Magnetic moment	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-25},10^{-23})$
		B	Magnetic flux density	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-3},10^{-1})$
		k	Boltzmann constant	V, F	C, F, P	$\mathcal{U}(1,5)$	1.381×10^{-23}
		T	Temperature	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{1},10^{3})$
II.36.38	$x = \frac{\mu H}{kT} + \frac{\mu\lambda}{\epsilon c^2 kT}M$	x	Parameter of magnetization	V, F	V, F	N/A	N/A
		μ	Magnetic moment	V, F	V, F	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-25},10^{-23})$
		H	Magnetic field strength	V, F	V, F	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-3},10^{-1})$
		k	Boltzmann constant	V, F	C, F, P	$\mathcal{U}(1,3)$	1.381×10^{-23}
		T	Temperature	V, F	V, F, P	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{1},10^{3})$
		λ	Constant	V, F	V, F, NN	$\mathcal{U}(1,3)$	$\mathcal{U}(0,1)$
		ϵ	Vacuum permittivity	V, F	C, F, P	$\mathcal{U}(1,3)$	8.854×10^{-12}
		c	Speed of light	V, F	C, F, P	$\mathcal{U}(1,3)$	2.998×10^{8}
		M	Number of magnetized atoms	V, F	V, I⋆, P	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{23},10^{25})$
III.4.33	$E = \frac{h\omega}{2\pi(\exp(h\omega/2\pi kT)-1)}$	E	Energy	V, F	V, F, P	N/A	N/A
		h	Planck constant	V, F	C, F, P	$\mathcal{U}(1,5)$	6.626×10^{-34}
		ω	Frequency	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{8},10^{10})$
		k	Boltzmann constant	V, F	C, F, P	$\mathcal{U}(1,5)$	1.381×10^{-23}
		T	Temperature	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{1},10^{3})$
III.9.52	$P_{\mathrm{I}\to\mathrm{II}} = \left(\frac{2\pi\mu E t}{h}\right)^2 \frac{\sin^2((\omega-\omega_0)t/2)}{(\omega-\omega_0)t/2)^2}$	$P_{\mathrm{I}\to\mathrm{II}}$	Probability	V, F	V, F, NN	N/A	N/A
		μ	Electric dipole moment	V, F	V, F	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-22},10^{-20})$
		E	Magnitude of electric field	V, F	V, F	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{1},10^{3})$
		t	Time	V, F	V, F, NN	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-18},10^{-16})$
		h	Planck constant	V, F	C, F, P	$\mathcal{U}(1,3)$	6.626×10^{-34}
		ω	Frequency	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{8},10^{10})$
		ω_0	Resonant frequency	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{8},10^{10})$

Table S10: Hard set of our proposed datasets (part 4).

Eq. ID	Formula	Symbols	Properties Original	Properties Ours	Distributions Original	Distributions Ours
III.10.19	$E = \mu\sqrt{B_x^2 + B_y^2 + B_z^2}$	E Energy	V, F	V, F	N/A	N/A
		μ Magnetic moment	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-25}, 10^{-23})$
		B_x Element of magnetic field	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$
		B_y Element of magnetic field	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$
		B_z Element of magnetic field	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$
III.21.20	$J = -\rho \frac{q}{m} A$	J Electric Current	V, F	V, F	N/A	N/A
		ρ Electric charge density	V, F	V, F, N	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{27}, 10^{29})$
		q Electric charge	V, F	V, F, N	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-11}, 10^{-9})$
		A Magnetic vector potential	V, F	V, F	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$
		m Mass	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{-30}, 10^{-28})$
B1	$A = \left(\frac{Z_1 Z_2 \alpha h c}{4E\sin^2(\theta/2)}\right)^2$	A Differential scattering cross section	V, F	V, F	N/A	N/A
		Z_1 Atomic number	V, F	V, I,P	$\mathcal{U}(1, 2)$	$\mathcal{U}_{\log}(10^{0}, 10^{1})$
		Z_2 Atomic number	V, F	V, I,P	$\mathcal{U}(1, 2)$	$\mathcal{U}_{\log}(10^{0}, 10^{1})$
		α Fine structure constant	V, F	C, F, P	$\mathcal{U}(1, 5)$	7.297×10^{-3}
		h Dirac's constant	V, F	C, F, P	$\mathcal{U}(1, 2)$	1.055×10^{-34}
		c Speed of light	V, F	C, F, P	$\mathcal{U}(1, 2)$	2.998×10^{8}
		E Non-relativistic kinetic energy	V, F	V, F, P	$\mathcal{U}(1, 3)$	$\mathcal{U}_{\log}(10^{-18}, 10^{-16})$
		θ Scattering angle	V, F	V, F, NN	$\mathcal{U}(1, 3)$	$\mathcal{U}(0, 2\pi)$
B2	$k = \frac{mk_G}{L^2}\left(1 + \sqrt{1 + \frac{2EL^2}{mk_G^2}}\cos(\theta_1 - \theta_2)\right)$	k Variable	V, F	V, F	N/A	N/A
		m Mass (The Earth)	V, F	V, F, P	$\mathcal{U}(1, 3)$	$\mathcal{U}_{\log}(10^{23}, 10^{25})$
		k_G Variable	V, F	V, F, P	$\mathcal{U}(1, 3)$	$\mathcal{U}_{\log}(10^{9}, 10^{11})$
		L Distance	V, F	V, F, P	$\mathcal{U}(1, 3)$	$\mathcal{U}_{\log}(10^{8}, 10^{10})$
		E Energy	V, F	V, F, P	$\mathcal{U}(1, 3)$	$\mathcal{U}_{\log}(10^{25}, 10^{27})$
		θ_1 Angle	V, F	V, F, NN	$\mathcal{U}(0, 6)$	$\mathcal{U}(0, 2\pi)$
		θ_2 Angle	V, F	V, F, NN	$\mathcal{U}(0, 6)$	$\mathcal{U}(0, 2\pi)$
B3	$r = \frac{d(1-\alpha^2)}{1+\alpha\cos(\theta_1 - \theta_2)}$	r Distance	V, F	V, F, P	N/A	N/A
		d Semimajor axis of elliptical orbit	V, F	V, F, P	$\mathcal{U}(1, 3)$	$\mathcal{U}_{\log}(10^{8}, 10^{10})$
		α Orbital eccentricity	V, F	V, F, P	$\mathcal{U}(2, 4)$	$\mathcal{U}(0, 1)$
		θ_1 Angle	V, F	V, F, NN	$\mathcal{U}(4, 5)$	$\mathcal{U}(0, 2\pi)$
		θ_2 Angle	V, F	V, F, NN	$\mathcal{U}(4, 5)$	$\mathcal{U}(0, 2\pi)$
B4	$v = \sqrt{\frac{2}{m}\left(E - U - \frac{L^2}{2mr^2}\right)}$	v Velocity	V, F	V, F, P	N/A	N/A
		m Mass (The Earth)	V, F	V, F, P	$\mathcal{U}(1, 3)$	$\mathcal{U}_{\log}(10^{23}, 10^{25})$
		E Energy	V, F	V, F, P	$\mathcal{U}(8, 12)$	$\mathcal{U}_{\log}(10^{25}, 10^{27})$
		U Potential energy	V, F	V, F, P	$\mathcal{U}(1, 3)$	$\mathcal{U}_{\log}(10^{25}, 10^{27})$
		L Angular momentum	V, F	V, F	$\mathcal{U}(1, 3)$	$\mathcal{U}_{\log}(10^{8}, 10^{10})$
		r Distance	V, F	V, F, P	$\mathcal{U}(1, 3)$	$\mathcal{U}_{\log}(10^{8}, 10^{10})$
B5	$t = \frac{2\pi d^{3/2}}{\sqrt{G(m_1+m_2)}}$	t Orbital period	V, F	V, F, P	N/A	N/A
		d Semimajor axis of elliptical orbit	V, F	V, F, P	$\mathcal{U}(1, 3)$	$\mathcal{U}_{\log}(10^{8}, 10^{10})$
		G Gravitational constant	V, F	C, F, P	$\mathcal{U}(1, 3)$	6.674×10^{-11}
		m_1 Mass (The Earth)	V, F	V, F, P	$\mathcal{U}(1, 3)$	$\mathcal{U}_{\log}(10^{23}, 10^{25})$
		m_2 Mass (The Earth)	V, F	V, F, P	$\mathcal{U}(1, 3)$	$\mathcal{U}_{\log}(10^{23}, 10^{25})$
B6	$\alpha = \sqrt{1 + \frac{2\epsilon^2 EL^2}{m(Z_1 Z_2 q^2)^2}}$	α Orbital eccentricity	V, F	V, F, P	N/A	N/A
		ϵ Energy	V, F	V, F	$\mathcal{U}(1, 3)$	$\mathcal{U}_{\log}(10^{-18}, 10^{-16})$
		E Energy	V, F	V, F, P	$\mathcal{U}(1, 3)$	$\mathcal{U}_{\log}(10^{-18}, 10^{-16})$
		L Distance	V, F	V, F, P	$\mathcal{U}(1, 3)$	$\mathcal{U}_{\log}(10^{-10}, 10^{-8})$
		m Mass	V, F	V, F, P	$\mathcal{U}(1, 3)$	$\mathcal{U}_{\log}(10^{-30}, 10^{-28})$
		Z_1 Atomic number	V, F	V, I,P	$\mathcal{U}(1, 3)$	$\mathcal{U}_{\log}(10^{0}, 10^{1})$
		Z_2 Atomic number	V, F	V, I,P	$\mathcal{U}(1, 3)$	$\mathcal{U}_{\log}(10^{0}, 10^{1})$
		q Electric charge	V, F	V, F	$\mathcal{U}(1, 3)$	$\mathcal{U}_{\log}(10^{-11}, 10^{-9})$
B7	$H = \sqrt{\frac{8\pi G\rho}{3} - \frac{k_f c^2}{a_f^2}}$	H Hubble's constant	V, F	V, F, P	N/A	N/A
		G Gravitational constant	V, F	C, F, P	$\mathcal{U}(1, 3)$	6.674×10^{-11}
		ρ Density of the Universe	V, F	V, F, P	$\mathcal{U}(1, 3)$	$\mathcal{U}_{\log}(10^{-27}, 10^{-25})$
		k_f Spacetime curvature	V, F	V, I	$\mathcal{U}(1, 2)$	$\mathcal{U}(-1, 1)$
		c Speed of light	V, F	C, F, P	$\mathcal{U}(1, 2)$	2.998×10^{8}
		a_f Radius	V, F	V, F, P	$\mathcal{U}(1, 3)$	$\mathcal{U}_{\log}(10^{8}, 10^{10})$
B9	$P = -\frac{32}{5}\frac{G^4}{c^5}\frac{(m_1 m_2)^2(m_1+m_2)}{r^5}$	P Gravitational wave energy	V, F	V, F	N/A	N/A
		G Gravitational constant	V, F	C, F, P	$\mathcal{U}(1, 2)$	6.674×10^{-11}
		c Speed of light	V, F	C, F, P	$\mathcal{U}(1, 2)$	2.998×10^{8}
		m_1 Mass	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{23}, 10^{25})$
		m_2 Mass	V, F	V, F, P	$\mathcal{U}(1, 5)$	$\mathcal{U}_{\log}(10^{23}, 10^{25})$
		r Distance	V, F	V, F, P	$\mathcal{U}(1, 2)$	$\mathcal{U}_{\log}(10^{8}, 10^{10})$

Table S11: Hard set of our proposed datasets (part 5).

Eq. ID	Formula	Symbols		Properties Original	Properties Ours	Distributions Original	Distributions Ours
B10	$\cos\theta_1 = \frac{\cos\theta_2 - v/c}{(1-v/c)\cos\theta_2}$	$\cos\theta_1$	Value	V, F	V, F	N/A	N/A
		θ_2	Angle	V, F	V, F	$\mathcal{U}(1,3)$	$\mathcal{U}(0,2\pi)$
		v	Velocity	V, F	V, F	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^5, 10^7)$
		c	Speed of light	V, F	C, F, P	$\mathcal{U}(4,6)$	2.998×10^8
B11	$I = I_0\left(\frac{\sin(\alpha/2)}{\alpha/2}\frac{\sin(N\delta/2)}{\sin(\delta/2)}\right)^2$	I	Wave intensity	V, F	V, F, P	N/A	N/A
		I_0	Amplitude of wave	V, F	V, F, P	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$
		α	Wavelength of X-ray	V, F	V, F, P	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-11}, 10^{-9})$
		N	Number of phase difference	V, F	V, I,P	$\mathcal{U}(1,2)$	$\mathcal{U}_{\log}(10^{0}, 10^{2})$
		δ	Wavelength of X-ray	V, F	V, F, P	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-11}, 10^{-9})$
B12	$F = \frac{q}{4\pi\epsilon y^2}\left(4\pi\epsilon V_e d - \frac{qdy^3}{(y^2-d^2)^2}\right)$	F	Force	V, F	V, F	N/A	N/A
		q	Electric charge	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$
		ϵ	Vacuum permittivity	V, F	C, F, P	$\mathcal{U}(1,5)$	8.854×10^{-12}
		y	Distance	V, F	V, F, P	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-2}, 10^{0})$
		V_e	Voltage	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		d	Distance	V, F	V, F, P	$\mathcal{U}(4,6)$	$\mathcal{U}_{\log}(10^{-2}, 10^{0})$
B13	$V_e = \frac{q}{4\pi\epsilon\sqrt{r^2+d^2-2dr\cos\alpha}}$	V_e	Potential	V, F	V, F	N/A	N/A
		ϵ	permittivity	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-12}, 10^{-10})$
		q	Electric charge	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-3}, 10^{-1})$
		r	Distance	V, F	V, F, P	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{-2}, 10^{0})$
		d	Distance between dipoles	V, F	V, F, P	$\mathcal{U}(4,6)$	$\mathcal{U}_{\log}(10^{-2}, 10^{0})$
		α	Angle	V, F	V, F	$\mathcal{U}(0,6)$	$\mathcal{U}(0,2\pi)$
B14	$V_e = E_f\cos\theta\left(\frac{\alpha-1}{\alpha+2}\frac{d^3}{r^2} - r\right)$	V_e	Potential (out)	V, F	V, F	N/A	N/A
		E_f	Magnitude of electric field	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{1}, 10^{3})$
		θ	Angle	V, F	V, F	$\mathcal{U}(0,6)$	$\mathcal{U}(0,2\pi)$
		r	Distance	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-2}, 10^{0})$
		d	Radius of dielectric sphere	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-2}, 10^{0})$
		α	Polarizability	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
B15	$\omega_0 = \frac{\sqrt{1-\frac{v^2}{c^2}}}{1+\frac{v}{c}\cos\theta}\omega$	ω_0	Frequency of electromagnetic waves	V, F	V, F	N/A	N/A
		v	Velocity	V, F	V, F, P	$\mathcal{U}(1,3)$	$\mathcal{U}_{\log}(10^{5}, 10^{7})$
		c	Speed of light	V, F	C, F, P	$\mathcal{U}(5,20)$	2.998×10^8
		ω	Frequency of electromagnetic waves	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{9}, 10^{11})$
		θ	Angle	V, F	V, F	$\mathcal{U}(0,6)$	$\mathcal{U}(0,2\pi)$
B16	$E = qV_e + \sqrt{(p-qA)^2c^2 + m^2c^4}$	E	Energy	V, F	V, F	N/A	N/A
		p	Momentum	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-9}, 10^{-7})$
		q	Electric charge	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-11}, 10^{-9})$
		A	Vector potential	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{1}, 10^{3})$
		c	Speed of light	V, F	C, F, P	$\mathcal{U}(1,5)$	2.998×10^8
		m	Mass	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-30}, 10^{-28})$
		V_e	Voltage	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
B17	$E = \frac{1}{2m}\left(p^2 + m^2\omega^2x^2\left(1+\alpha\frac{x}{y}\right)\right)$	E	Energy	V, F	V, F	N/A	N/A
		m	Mass	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-30}, 10^{-28})$
		p	Momentum	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-9}, 10^{-7})$
		ω	Frequency of electromagnetic waves	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{9}, 10^{11})$
		x	Position	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-11}, 10^{-9})$
		α	Deviation from the harmonic oscillator	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-1}, 10^{1})$
		y	Distance	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{-11}, 10^{-9})$
B19	$p_f = -\frac{1}{8\pi G}\left(\frac{c^4k_f}{a_f^2} + c^2H^2(1-2\alpha)\right)$	p_f	Pressure	V, F	V, F	N/A	N/A
		G	Gravitational constant	V, F	C, F, P	$\mathcal{U}(1,5)$	6.674×10^{-11}
		c	Speed of light	V, F	C, F, P	$\mathcal{U}(1,5)$	2.998×10^8
		k_f	Variable	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{1}, 10^{3})$
		a_f	Distance	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{8}, 10^{10})$
		H	Variable	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{0}, 10^{2})$
		α	Variable	V, F	V, F	$\mathcal{U}(1,5)$	$\mathcal{U}(-10,10)$
B20	$A = \frac{\alpha^2h^2}{4\pi m^2c^2}\left(\frac{\omega_0}{\omega}\right)^2\left(\frac{\omega_0}{\omega} + \frac{\omega}{\omega_0} - \sin^2\theta\right)$	A	Differential cross section	V, F	V, F	N/A	N/A
		α	Fine structure constant	V, F	C, F, P	$\mathcal{U}(1,5)$	7.297×10^{-3}
		h	Planck constant	V, F	C, F, P	$\mathcal{U}(1,5)$	6.626×10^{-34}
		m	Electron mass	V, F	C, F, P	$\mathcal{U}(1,5)$	9.109×10^{-31}
		c	Speed of light	V, F	C, F, P	$\mathcal{U}(1,5)$	2.998×10^8
		ω_0	Frequency	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{9}, 10^{11})$
		ω	Frequency	V, F	V, F, P	$\mathcal{U}(1,5)$	$\mathcal{U}_{\log}(10^{9}, 10^{11})$
		θ	Scattering angle	V, F	V, F	$\mathcal{U}(0,6)$	$\mathcal{U}(0,2\pi)$

B Hyperparameters for Five Existing SR Baselines

Table S12 shows the hyperparameter space for the five existing symbolic regression baselines. The hyperparameters of gplearn [3] [11], AFP [38], and AFP-FE [35] [12] are optimized by Optuna [40], a hyperparameter optimization framework.

Table S12: Hyperparameter sets for the five existing symbolic regression baselines.

Method	Hyperparameter sets
gplearn	100 trials with random combinations of the following hyperparameter spaces: *population_size*: $\mathcal{U}(10^2, 10^3)$, *generations*: $\mathcal{U}(10, 10^2)$, *stopping_criteria*: $\mathcal{U}(10^{-10}, 10^{-2})$, *warm_start*: {True, False}, *const_range*: {None, $(-1.0, 1.0), (-10, 10), (-10^2, 10^2), (-10^3, 10^3), (-10^4, 10^4)$}, *max_samples*: $\mathcal{U}(0.9, 1.0)$, *parsimony_coefficient*: $\mathcal{U}(10^{-3}, 10^{-2})$
AFP	100 trials with random combinations of the following hyperparameter spaces: *popsize*: $\mathcal{U}(100, 1000)$, *g*: $\mathcal{U}(250, 2500)$, *stop_threshold*: $\mathcal{U}(10^{-10}, 10^{-2})$, *op_list*: {['n', 'v', '+', '-', '*', '/', 'exp', 'log', '2', '3', 'sqrt'], ['n', 'v', '+', '-', '*', '/', 'exp', 'log', '2', '3', 'sqrt', 'sin', 'cos']}
AFP-FE	100 trials with random combinations of the following hyperparameter spaces: *popsize*: $\mathcal{U}(100, 1000)$, *g*: $\mathcal{U}(250, 2500)$, *stop_threshold*: $\mathcal{U}(10^{-10}, 10^{-2})$, *op_list*: {['n', 'v', '+', '-', '*', '/', 'exp', 'log', '2', '3', 'sqrt'], ['n', 'v', '+', '-', '*', '/', 'exp', 'log', '2', '3', 'sqrt', 'sin', 'cos']}
AI Feynman	{*bftt*: 60, *epoch*: 300, *op*: '7ops.txt', *poly_deg*: 3}, {*bftt*: 60, *epoch*: 300, *op*: '10ops.txt', *poly_deg*: 3}, {*bftt*: 60, *epoch*: 300, *op*: '14ops.txt', *poly_deg*: 3}, {*bftt*: 60, *epoch*: 300, *op*: '19ops.txt', *poly_deg*: 3}, {*bftt*: 120, *epoch*: 300, *op*: '14ops.txt', *poly_deg*: 4}, {*bftt*: 120, *epoch*: 300, *op*: '19ops.txt', *poly_deg*: 4}, {*bftt*: 60, *epoch*: 500, *op*: '7ops.txt', *poly_deg*: 3}, {*bftt*: 60, *epoch*: 500, *op*: '10ops.txt', *poly_deg*: 3}, {*bftt*: 60, *epoch*: 500, *op*: '14ops.txt', *poly_deg*: 3}, {*bftt*: 60, *epoch*: 500, *op*: '19ops.txt', *poly_deg*: 3}
DSR	{*seed*: 1, *function_set*: ['add', 'sub', 'mul', 'div', 'sin', 'cos', 'exp', 'log'}, {*seed*: 2, *function_set*: ['add', 'sub', 'mul', 'div', 'sin', 'cos', 'exp', 'log'}, {*seed*: 3, *function_set*: ['add', 'sub', 'mul', 'div', 'sin', 'cos', 'exp', 'log'}, {*seed*: 4, *function_set*: ['add', 'sub', 'mul', 'div', 'sin', 'cos', 'exp', 'log'}, {*seed*: 5, *function_set*: ['add', 'sub', 'mul', 'div', 'sin', 'cos', 'exp', 'log'}, {*seed*: 1, *function_set*: ['add', 'sub', 'mul', 'div', 'sin', 'cos', 'exp', 'log', 'const']}, {*seed*: 2, *function_set*: ['add', 'sub', 'mul', 'div', 'sin', 'cos', 'exp', 'log', 'const']}, {*seed*: 3, *function_set*: ['add', 'sub', 'mul', 'div', 'sin', 'cos', 'exp', 'log', 'const']}, {*seed*: 4, *function_set*: ['add', 'sub', 'mul', 'div', 'sin', 'cos', 'exp', 'log', 'const']}, {*seed*: 5, *function_set*: ['add', 'sub', 'mul', 'div', 'sin', 'cos', 'exp', 'log', 'const']}

[11] `https://gplearn.readthedocs.io/en/stable/reference.html#symbolic-regressor`
[12] `https://github.com/cavalab/ellyn`