Does LLM dream of differential equation discovery?

Elizaveta Ivanchik, Timur Bavshin, Alexander Hvatov*

NSS Lab, ITMO University, Saint Petersburg, Russia

Introduction

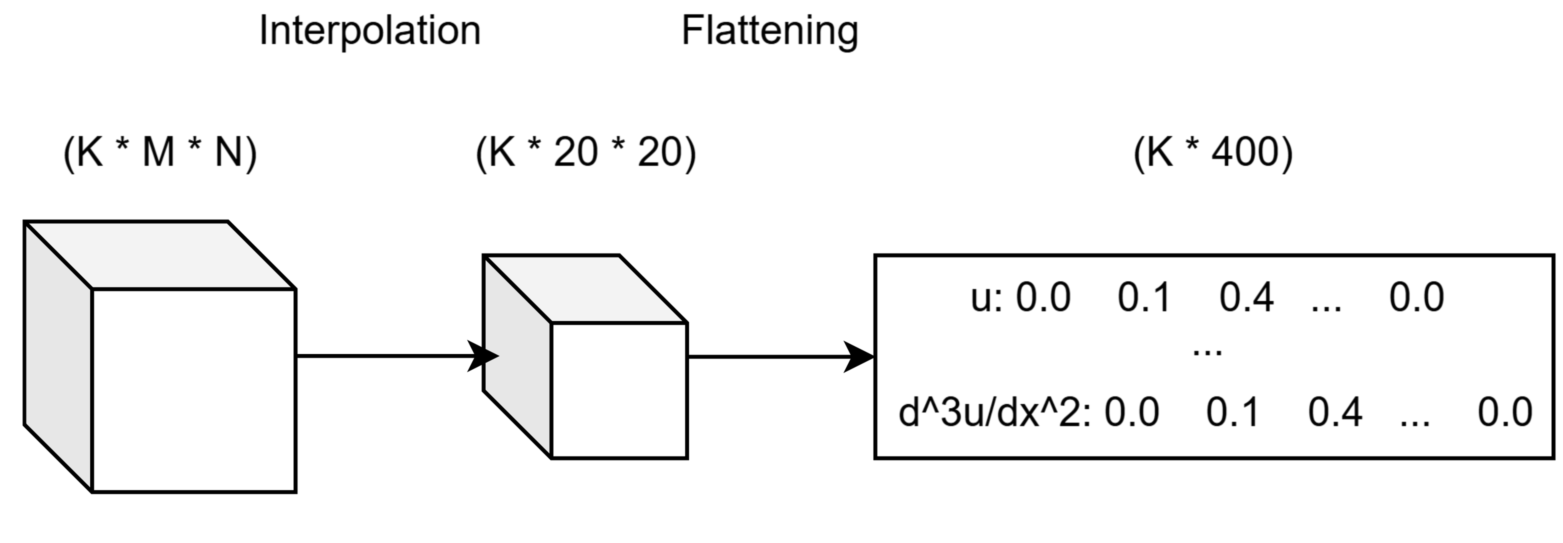
The challenge of PDE discovery: Discovering the underlying Partial Differential Equations (PDEs) from observed physical data is a core task in scientific machine learning. It requires handling differential operators, which LLMs are not natively equipped to process from data.

The LLM opportunity: Large Language Models (LLMs) possess vast scientific knowledge from textbooks but struggle to connect this knowledge to raw, numerical physical field data. Can we bridge this gap?

Our goal: This work investigates if general-purpose LLMs, without specific fine-tuning, can effectively contribute to the PDE discovery process when physical data and the task are formatted appropriately.

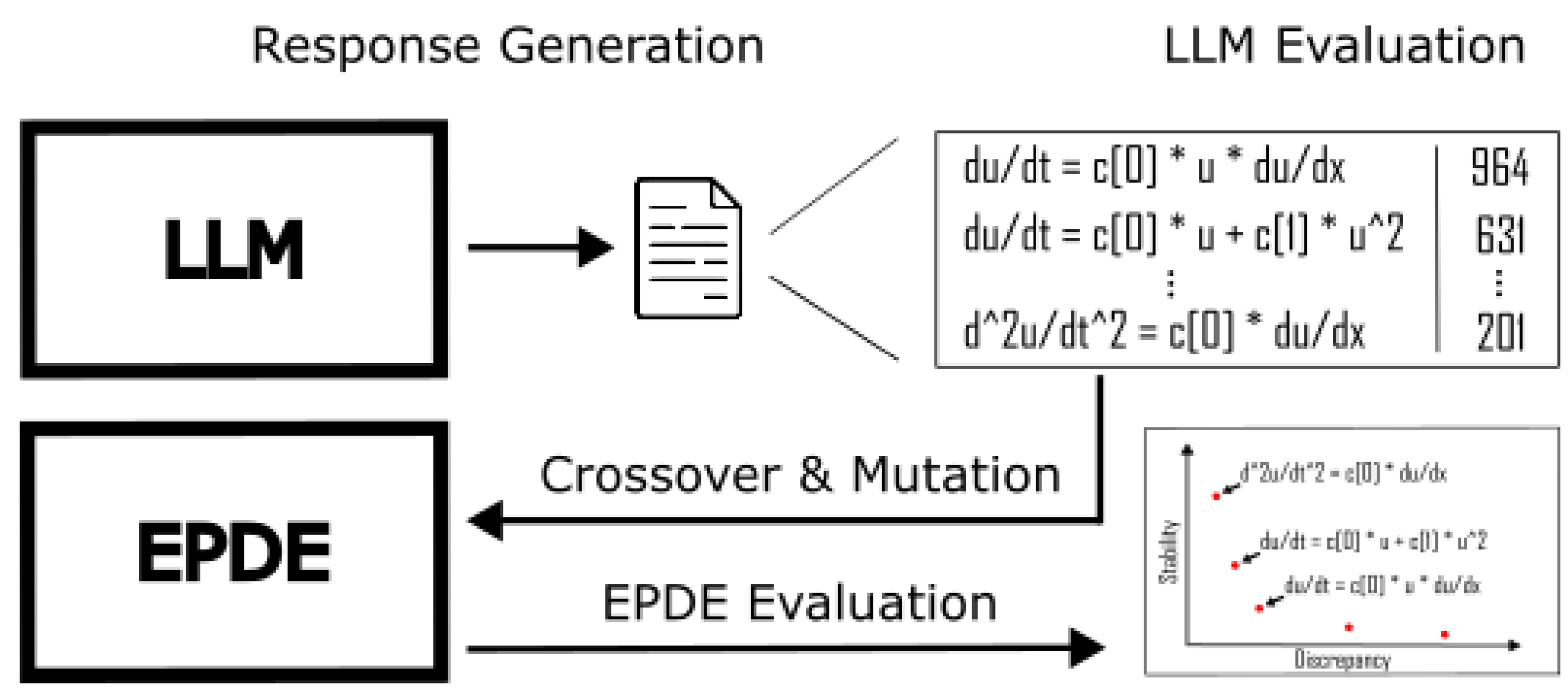
Proposed Approach

Our methodology formulates PDE discovery as a **code-generation task** for LLMs and integrates them into a **meta-learning loop** with an evolutionary algorithm, EPDE. The process begins with preparing **the physical field data** for the LLM. Then data is **downsampled** to a coarse grid to fit the model's context window while preserving essential physics.



K -- data field and derivatives, M and N -- amount of X and T points in the grid

The core of our approach uses the **LLM as a hypothesis generator**. A carefully engineered prompt guides the model, containing instructions, a code snippet for evaluation, the formatted input data, and a critical experience buffer that records the performance of previously proposed equations. The LLM then **generates Python code** defining a **candidate equation**. This output is extracted, post-processed for validity, and evaluated.



EPDE then refines these **candidates** using its **mutation and crossover** operators to perform a precise numerical optimization, converging on the final equation form and coefficients.

Experiments

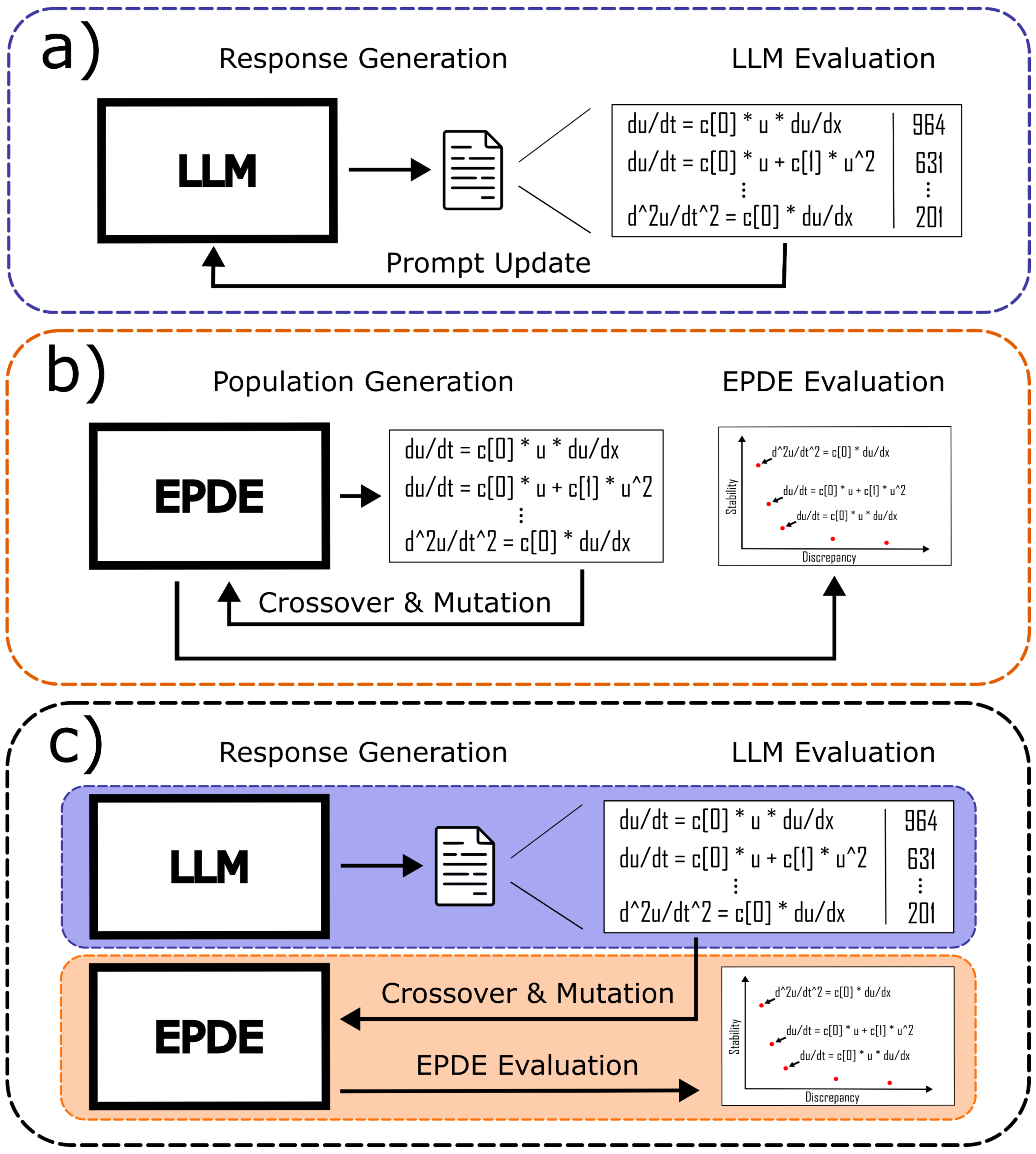
We evaluated our approach on three canonical PDEs — Burgers', Wave, and Korteweg–de Vries — using both clean and noisy data. Key metrics were the **discovery rate**, and the **coefficient error** (10^{-4}).

Dataset	EPDE		LLM		EPDE+LLM	
	DR	CE	DR	CE	DR	CE
Wave	0.97	**7.54**	0.97	657	**1.00**	**7.54**
Burgers A	0.53	**0.85**	0.86	3.94	**0.90**	**0.85**
Burgers B	0.50	**4.55**	0.53	90.5	**0.90**	**4.55**
KdV	0.10	**154**	0.13	192	**0.37**	**154**

Clear data performance comparison

Dataset	EPDE		LLM		EPDE+LLM	
	DR	CE	DR	CE	DR	CE
Wave	0.03	998	0.07	2546	**0.20**	**18.2**
Burgers A	0.03	576	**0.80**	**86.1**	0.23	376
Burgers B	0.03	**858**	0.07	4967	**0.20**	1206
KdV	0.03	**262**	0.00	-	**0.30**	291

Noisy data performance comparison



Acknowledgement: This work supported by the Ministry of Economic Development of the Russian Federation (IGK 000000C313925P4C0002), agreement No139-15-2025-010

*alex_hvatov@itmo.ru