NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Teaching LLMs Mathematical Reasoning for Wireless Communications

WirelessMathBench & WirelessMathLM

Xin Li

Nanyang Technological University, Singapore

xin019@e.ntu.edu.sg https://lixin.ai

October 27, 2025

Outline

The Promise: AI as Your Engineering Assistant

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

The Vision

LLMs have mastered **coding, writing, reasoning** — why not **engineering problems**?

Imagine: An AI that can design wireless systems for you

What We Want

Natural Language Input:

- "Design a 5G beamforming system for 8 users with 64 antennas"
- "Optimize RIS phase shifts to maximize sum rate"
- "Derive the capacity region for MIMO interference channel"

Expected Output

Complete Solution:

- Mathematical model formulation
- Step-by-step derivations
- Optimized parameter values
- Performance guarantees

Mathematical Modeling: The Core of Wireless

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

The Foundation of Our Field

Physical World $\Rightarrow$ **Mathematical Models** $\Rightarrow$ System Optimization

Typical Problem Structure

1. **System Modeling:**
   - Channel: $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$
2. **Performance Metric:**
   - Capacity: $C = \log_2(1 + \mathrm{SNR})$
   - SINR: $\gamma = \dfrac{|\mathbf{h}^{\mathrm{H}}\mathbf{w}|^2}{\sum_{i \neq k} |\mathbf{h}_i^{\mathrm{H}}\mathbf{w}|^2 + \sigma^2}$

Challenges for LLMs

- **Complex-valued ops:** $\mathbb{C}^{M \times N}$ matrices
- **Matrix calculus:** $\nabla_{\mathbf{W}} \mathrm{Tr}(\mathbf{W}\mathbf{A})$
- **Physical constraints:** power, causality, stability
- **Multi-step reasoning:** model $\rightarrow$ simplify $\rightarrow$ optimize

Can LLMs master this? That's what we investigate.

Roadmap for Today's Talk

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

1. **WirelessMathBench:** Dataset construction, problem types, evaluation protocol
2. **Baseline Evaluation:** How good are today's LLMs? (GPT, Deepseek, Qwen, etc.)
3. **WirelessMathLM:** Data, GRPO training method, implementation
4. **Results:** Performance gains, ablation studies, error analysis
5. **Discussion:** Limitations, future directions, broader impact

Let's dive into the technical details!

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Part 1

WirelessMathBench:

A Mathematical Modeling Benchmark for LLMs in Wireless Communications

Xin Li, Mengbing Liu, Li Wei, Jiancheng An, Mérouane Debbah, Chau Yuen

ACL Findings 2025

Existing Math Benchmarks: The Landscape

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Benchmark	Difficulty	Domain	Engineering	Size
GSM8K	Elementary	General	No	1,319
MATH	High School	General	No	5,000
OCWCourses	University	General	No	272
MMMU	University	Multi	Partial	1,983
OlympiadBench	Competition	General	No	8,476
SciBench	University	Science	Partial	695
Ours	**Expert/Research**	**Wireless**	**Yes**	**587**

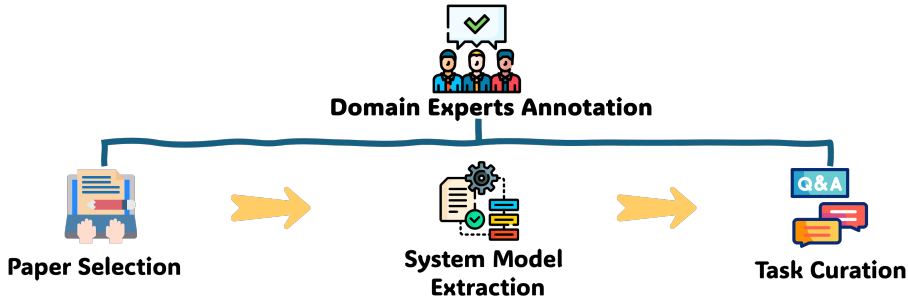
Gap in Existing Work

- Focus on general mathematics
- Limited technical domains
- Lack engineering constraints
- Missing real-world complexity

Our Contribution

- Expert/research level
- Real engineering problems
- Complete system models
- Verifiable correctness

Dataset Construction: Four-Stage Pipeline



Rigorous Quality Control

Stage 1: Paper Selection (IEEE TWC, JSAC, TCOM, ICC, GLOBECOM)

Stage 2: System Model Extraction (LLM-assisted + manual refinement)

Stage 3: Question Generation (MCQ + Progressive + FEC)

Across Stages: Expert Validation (5 researchers, multi-round review)

Stage 1: Paper Selection & Coverage

Selection Criteria

Source Materials:

- **40 state-of-the-art papers**
- Top-tier venues
- Freely accessible on **arXiv**

Content Requirements:

- **Nontrivial mathematical derivations**
- Physical & dimensional constraints

Keyword Distribution

Technical Coverage: Detailed Distribution

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Model-Based Topics

System Model	#
RIS (Reconfigurable Int. Surfaces)	19
MIMO / Massive MIMO	12
UAV Communications	6
ISAC (Integrated Sensing & Comm.)	6
Satellite Communications	4
SIM (Stacked Int. Metasurface)	3
NOMA (Non-Orthog. Multiple Access)	2

Problem-Based Topics

Problem Domain	#
Beamforming Design	18
Channel Estimation	12
Performance Analysis	8
Trajectory Design	5
Power Allocation	5
Resource Management	4

Note

Papers may span multiple categories · Total unique papers: 40

Stage 2: System Model Extraction

Unified Summarize Extraction

1. Automated Parsing

- LaTeX source analysis
- LLM-assisted Equation identification & Variable & Context extraction

2. Manual Refinement

- Verify completeness
- Add missing definitions
- Clarify notation
- Ensure self-containment

Example: RIS System

Extracted Components:

System: RIS-assisted downlink

Channel: $\mathbf{h}_{\text{eff}} = \mathbf{h}_{\text{d}} + \mathbf{H}_{\text{r}}^{H} \hat{} \mathbf{h}_{\text{t}}$

Variables:

- $\mathbf{h}_{\text{d}} \in \mathbb{C}^{M \times 1}$: direct channel
- $\mathbf{H}_{\text{r}} \in \mathbb{C}^{N \times M}$: BS-RIS channel
- $\mathbf{h}_{\text{t}} \in \mathbb{C}^{N \times 1}$: RIS-user channel
- $\hat{} = \text{diag}(e^{j\theta_1}, ..., e^{j\theta_N})$

Constraints: $|\theta_n| = 1, \forall n$

Anti-Contamination

Reformulate in original language to avoid word-for-word reproduction

Stage 3: Question Design Philosophy

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Three Task Types with Progressive Difficulty

MCQ

Multiple Choice

Test recognition and recall

Example:

Which expression gives MRC combining gain?

A) $\frac{|h_1|^2}{\sigma^2}$

B) $\frac{\sum |h_i|^2}{\sigma^2}$ ✓

C) $\frac{\max |h_i|^2}{\sigma^2}$

D) $\frac{N|h_1|^2}{\sigma^2}$

Fill-in

Progressive Masking

3 difficulty levels

Level 1 (25%):

$\gamma = \frac{[\text{MASK}]}{\sigma^2}$

Level 2 (50%):

$\gamma = \frac{\sum [\text{M1}]}{[\text{M2}]}$

Level 3 (75%):

$\gamma = \frac{[\text{M1}]}{[\text{M2}]} = \frac{|\mathbf{h}^H[\text{M3}]|^2}{[\text{M4}]}$

FEC

Full Equation

Complete derivation

Example:

Given: N-antenna MRC receiver, channel $\mathbf{h}$*, noise* σ^2

Derive: Complete SINR expression with optimal weights

Question Example: Multiple Choice Question (MCQ)

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Design Philosophy

Purpose:

- Test recognition and recall ability
- Evaluate understanding of key wireless system modeling elements

Key Challenge

Models must distinguish between closely related expressions that differ only in critical details (operators, dimensions, sequences)

Background

In a double-RIS-assisted massive MIMO system, the overall channel from user k, denoted $\boldsymbol{h}_k$, incorporates single-reflection links via RIS_1 and RIS_2, as well as a double-reflection link through both RISs. Here, $\boldsymbol{N}_j$ represents the channel from RIS_j to the BS, and $\boldsymbol{D}$ denotes the channel between the RISs. The reflection coefficients of RIS_j are given by $\boldsymbol{\theta}_j$.

Multiple Choice Question

$$\boldsymbol{h}_k = [MASK]$$

Question: Which expression correctly represents the user-k effective channel at the base station?

Options:

(A) $\boldsymbol{N}_2 \operatorname{diag}(\boldsymbol{h}_{k2})\, \boldsymbol{\theta}_2 + \boldsymbol{N}_1 \operatorname{diag}(\boldsymbol{h}_{k1})\, \boldsymbol{\theta}_1 + \boldsymbol{N}_2 \operatorname{diag}(\boldsymbol{\theta}_2)\, \boldsymbol{D} \operatorname{diag}(\boldsymbol{\theta}_1)\, \boldsymbol{h}_{k1}$.

(B) $\boldsymbol{N}_2 \boldsymbol{h}_{k2} \boldsymbol{\theta}_2 + \boldsymbol{N}_1 \operatorname{diag}(\boldsymbol{h}_{k1})\, \boldsymbol{\theta}_1 + \boldsymbol{D} \operatorname{diag}(\boldsymbol{\theta}_2)\, \boldsymbol{\theta}_1$

(C) $\boldsymbol{N}_2 \operatorname{diag}(\boldsymbol{h}_{k2})\, \boldsymbol{\theta}_2 + \boldsymbol{N}_1 \boldsymbol{h}_{k1} \boldsymbol{\theta}_1 + \boldsymbol{N}_2 \operatorname{diag}(\boldsymbol{\theta}_2)\, \boldsymbol{D} \operatorname{diag}(\boldsymbol{\theta}_1)\, \boldsymbol{h}_{k1}$

(D) $\boldsymbol{N}_2 \operatorname{diag}(\boldsymbol{h}_{k2})\, \boldsymbol{\theta}_2 + \boldsymbol{N}_1 \operatorname{diag}(\boldsymbol{h}_{k1})\, \boldsymbol{\theta}_1$

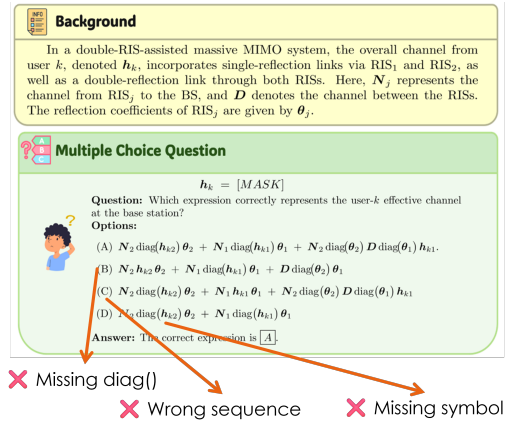
Answer: The correct expression is $\boxed{A}$.

MCQ Distractor Design: Error Analysis

Distractor Design Principles

Three Categories of Common Errors:

1. **Missing Operators** (Option B)
   - Omitting `diag()` operations
   - Tests understanding of matrix
2. **Wrong Sequence** (Option C)
   - Incorrect order of operations
   - Tests knowledge of signal flow
3. **Missing Symbols** (Option D)
   - Incomplete terms in expressions
   - Tests completeness of derivation



Design Goal

Each distractor represents a **plausible but incorrect** reasoning path models might follow

Question Example: Progressive Masking Fill-in-the-Blank

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Progressive Difficulty Design

Core Concept:

- Incremental complexity across 3 levels
- Each level is independent

Level 1 (25% masked):

- Single variable substitution
- Straightforward inference

Level 2 (50% masked):

- Two interdependent variables
- Requires understanding relationships

Level 3 (75% masked):

- Multiple structured terms
- Complex derivation needed

Background

In a double-RIS-assisted massive MIMO system, the overall channel from user k, denoted $\boldsymbol{h}_k$, incorporates single-reflection links via RIS_1 and RIS_2, as well as a double-reflection link through both RISs. Here, $\boldsymbol{N}_j$ represents the channel from RIS_j to the BS, and $\boldsymbol{D}$ denotes the channel between the RISs. The reflection coefficients of RIS_j are given by $\boldsymbol{\theta}_j$.

Progressive Masking Fill-in-the-blank (Level 1, 2, 3)

L1 $\boldsymbol{h}_k = \boldsymbol{N}_2 \operatorname{diag}(\boldsymbol{h}_{k2})\, [MASK] + \boldsymbol{N}_1 \operatorname{diag}(\boldsymbol{h}_{k1})\, \boldsymbol{\theta}_1 + \boldsymbol{N}_2 \operatorname{diag}(\boldsymbol{\theta}_2)\, \boldsymbol{D} \operatorname{diag}(\boldsymbol{\theta}_1)\, \boldsymbol{h}_{k1}$

Question 1: Which reflection vector is missing in the first single-reflection term?

L2 $\boldsymbol{h}_k = [MASK] \operatorname{diag}(\boldsymbol{h}_{k2})\, [MASK] + \boldsymbol{N}_1 \operatorname{diag}(\boldsymbol{h}_{k1})\, \boldsymbol{\theta}_1 + \boldsymbol{N}_2 \operatorname{diag}(\boldsymbol{\theta}_2)\, \boldsymbol{D} \operatorname{diag}(\boldsymbol{\theta}_1)\, \boldsymbol{h}_{k1}$

Question 2: Fill in the two interdependent channel-related variables for the RIS_2 path.

L3 $\boldsymbol{h}_k = [MASK] \operatorname{diag}(\boldsymbol{h}_{k2})\, [MASK] + [MASK] \operatorname{diag}(\boldsymbol{h}_{k1})\, \boldsymbol{\theta}_1 + \boldsymbol{N}_2 \operatorname{diag}(\boldsymbol{\theta}_2)\, \boldsymbol{D} \operatorname{diag}(\boldsymbol{\theta}_1)\, \boldsymbol{h}_{k1}$

Question 3: Complete the three missing terms describing the single-reflection link via RIS_1.

Mask Ratio **Answer**: $\boldsymbol{h}_k = \boldsymbol{N}_2 \operatorname{diag}(\boldsymbol{h}_{k2})\, \boldsymbol{\theta}_2 + \boldsymbol{N}_1 \operatorname{diag}(\boldsymbol{h}_{k1})\, \boldsymbol{\theta}_1 + \boldsymbol{N}_2 \operatorname{diag}(\boldsymbol{\theta}_2)\, \boldsymbol{D} \operatorname{diag}(\boldsymbol{\theta}_1)\, \boldsymbol{h}_{k1}$

Mask Ratio Interpretation

Higher masking $\rightarrow$ Less context provided $\rightarrow$ Greater reconstruction challenge

Question Example: Full Equation Completion (FEC)

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Maximum Difficulty Challenge

Task Description:

- **Complete formula hidden**
- Only scenario description provided

Required Capabilities:

- Deep domain knowledge
- Multi-step symbolic derivation
- Physical constraint awareness
- Dimensional consistency verification

Background

In a double-RIS-assisted massive MIMO system, the overall channel from user k, denoted $\boldsymbol{h}_k$, incorporates single-reflection links via RIS_1 and RIS_2, as well as a double-reflection link through both RISs. Here, $\boldsymbol{N}_j$ represents the channel from RIS_j to the BS, and $\boldsymbol{D}$ denotes the channel between the RISs. The reflection coefficients of RIS_j are given by $\boldsymbol{\theta}_j$.

Full Equation Completion

$$\boldsymbol{h}_k = [MASK]$$

Question: Write the full expression for the overall effective channel.

Answer: $\boldsymbol{h}_k = \boldsymbol{N}_2 \operatorname{diag}(\boldsymbol{h}_{k2})\, \boldsymbol{\theta}_2 + \boldsymbol{N}_1 \operatorname{diag}(\boldsymbol{h}_{k1})\, \boldsymbol{\theta}_1 + \boldsymbol{N}_2 \operatorname{diag}(\boldsymbol{\theta}_2)\, \boldsymbol{D} \operatorname{diag}(\boldsymbol{\theta}_1)\, \boldsymbol{h}_{k1}$

Expert-Level Performance

FEC represents the reasoning level expected from **human wireless engineers**

Expert Validation Process

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Review Protocol

1. Independent Review

- Each question reviewed by 2+ experts
- Feedback on accuracy and clarity
- Focus on technical correctness

2. Consensus Discussion

- Resolve disagreements collaboratively
- Third expert consultation if needed
- Iterate until full consensus reached

Validation Criteria

Technical Accuracy:

- Correct mathematical expressions
- Proper physical constraints
- Dimensional consistency

Question Quality:

- Clear problem statement
- Unambiguous correct answer
- No conflicting interpretations

Final Outcome

587 high-quality problems passed rigorous validation $\rightarrow$ Ready for benchmark evaluation

Dataset Statistics: Final Composition

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Question Type Distribution

Type	Count	%
MCQ	125	21.3%
Fill-in Level 1	120	20.4%
Fill-in Level 2	115	19.6%
Fill-in Level 3	112	19.1%
FEC	115	19.6%
Total	**587**	**100%**

Key Characteristics

- From **40** state-of-the-art research papers
- **Expert-level** difficulty & **Real-engineering** tasks

System Models (7 categories)

- RIS (19 papers) MIMO (12)
- UAV (6) ISAC (6)
- Satellite (4) SIM (3) NOMA (2)

Problem Domains (6 areas)

- Beamforming (18) Channel Est. (12)
- Performance Analysis (8)
- Trajectory Design (5) Power Alloc. (5)
- Resource Management (4)

Evaluation Protocol

Tested Models (16 Total)

Reasoning Models:

- DeepSeek-R1 (671B)
- OpenAI-o1 / o1-mini

General LLMs:

- GPT-4o, GPT-4
- DeepSeek-V3 (671B)
- Gemini-2.0-flash, 1.5-pro/flash

Math-Specialized:

- Qwen2.5-Math-72B/7B

Domain-Tuned:

- LLaMA-3-8B-Tele

Setup

Prompting:

- Zero-shot evaluation
- Standardized prompts

Scoring:

- MCQ: Direct comparison
- Fill-in and FEC: GPT-4o judge

Fair Comparison

- Identical prompts & Same evaluation
- No cherry-picking
- Reproducible setup

Main Results: The Performance Cliff

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Model	MCQ	Level 1	Level 2	Level 3	FEC	Avg
Reasoning Models						
DeepSeek-R1	76.0%	**60.0%**	**34.9%**	**12.5%**	**7.8%**	**38.1%**
OpenAI-o1	66.4%	59.2%	32.2%	8.0%	7.0%	34.6%
OpenAI-o1-mini	66.4%	53.3%	29.6%	10.7%	4.4%	32.9%
General Large Models						
GPT-4o	72.8%	42.5%	28.7%	6.3%	4.4%	30.9%
DeepSeek-V3	**78.4%**	50.0%	24.4%	6.3%	7.0%	33.2%
Gemini-2.0-flash	71.2%	40.8%	24.4%	5.4%	4.4%	29.2%
Math-Specialized Models						
Qwen2.5-Math-72B	70.4%	37.5%	26.1%	7.1%	6.1%	29.4%
Smaller/Domain Models						
Qwen2.5-Math-7B	58.4%	21.7%	7.0%	4.5%	1.7%	18.8%
LLaMA-3-8B-Tele	40.8%	11.7%	4.4%	2.7%	0.9%	12.1%

Model Comparison: Key Observations

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

1. Reasoning Advantage

DeepSeek-R1 vs DeepSeek-V3:

- -2.4 pts MCQ
- +10.0 pts Level 1
- +10.5 pts Level 2
- +6.2 pts Level 3
- +0.8 pts FEC

Explicit reasoning helps on complex derivations

2. Math Specialization

Qwen2.5-Math-72B vs LLaMA-70B:

- +4.8 pts average
- Similar on MCQ & Better on Fill-in

General math training provides limited help

3. Domain Fine-tuning

LLaMA-3-8B-Tele vs Base:

- -4.8 pts MCQ
- +0.9 pts Level 1
- -3.4 pts Level 2

Protocol training $\neq$ math reasoning

4. Scale Effects

Math-7B vs Math-72B:

- -12 pts MCQ
- -15.8 pts Level 1
- -19.1 pts Level 2

Scale matters, but even 72B models struggle (29.4%)

Error Analysis: Distribution & Insights

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Error Distribution

40 random failure samples from DeepSeek-R1:

- **Partial Fill Mismatch:** 31%
  - Merging separate placeholders
  - Inconsistent interdependent variables
- **Symbol Misinterpretation:** 29%
  - Wrong symbols (^{H} vs. ^{T})
  - Omitted key operators
- **Incorrect Derivation:** 24%
  - Missing intermediate steps
  - Error propagation
- **System Mixing:** 11%
  - Extraneous terms from other systems
- **Other:** 5%

Key Insights

Common Error Patterns:

- 60% errors involve **symbol-mistakes**
- Early mistakes propagatd
- Models struggle with domain-constraints

Implication

Current LLMs lack robust mechanisms for maintaining **multi-step symbolic consistency** in specialized engineering contexts

Error Example 1: Partial Fill Mismatch

Question: Cell-Free MIMO Conjugate Beamforming

Fill in the blanks:

$$s_m = [\text{MASK1}] \sum_{k=1}^{K} [\text{MASK2}] u_k$$

Given: P_m (power), η_{mk} (control coef.), $\hat{g}_{mk}$ (channel), u_k (symbol)

Correct Answer

$$s_m = \sqrt{P_m} \sum_{k=1}^{K} \sqrt{\eta_{mk}} \hat{g}_{mk}^* u_k$$

Ground Truth:

- MASK1 $= \sqrt{P_m}$
- MASK2 $= \sqrt{\eta_{mk}} \hat{g}_{mk}^*$

DeepSeek-R1 Output

$$s_m = \sqrt{P_m \eta_{mk}} \sum_{k=1}^{K} \hat{g}_{mk}^* u_k$$

Errors:

- Merged two masks into one: $\sqrt{P_m \eta_{mk}}$
- Moved η_{mk} outside summation

Error Example 2: Symbol Misinterpretation

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Question: RIS Channel Model

Complete the cascaded channel expression:

$$\mathbf{h}_{\text{eff}} = \mathbf{h}_{\text{d}} + [\text{MASK}]\hat{}\,\mathbf{h}_{\text{t}}$$

Given: $\mathbf{h}_{\text{d}}$ (direct link), $\mathbf{H}_{\text{r}}$ (BS-RIS, $\mathbb{C}^{N\times M}$), $\mathbf{h}_{\text{t}}$ (RIS-user, $\mathbb{C}^{N\times 1}$), $\hat{}$ (phase shift)

Correct Answer

$$\mathbf{h}_{\text{eff}} = \mathbf{h}_{\text{d}} + \mathbf{H}_{\text{r}}^{H}\hat{}\,\mathbf{h}_{\text{t}}$$

Ground Truth:

- MASK $= \mathbf{H}_{\text{r}}^{H}$
- Hermitian transpose for complex matrices

DeepSeek-R1 Output

$$\mathbf{h}_{\text{eff}} = \mathbf{h}_{\text{d}} + \mathbf{H}_{\text{r}}^{T}\hat{}\,\mathbf{h}_{\text{t}}$$

Errors:

- Used regular transpose ^{T} instead of Hermitian ^{H}
- Omits conjugation operation

Error Example 3: Incorrect Equation Derivation

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Question: MIMO Received Signal Power

Derive the received signal power at user k:

$$P_k^{\text{recv}} = [\text{MASK}]$$

Given: Transmit power ρ_k, channel $\mathbf{h}_k \in \mathbb{C}^{M\times 1}$, beamforming $\mathbf{w}_k \in \mathbb{C}^{M\times 1}$, $\|\mathbf{w}_k\|^2 = 1$

Correct Answer

$$P_k^{\text{recv}} = \rho_k |\mathbf{h}_k^H \mathbf{w}_k|^2$$

Derivation:

- Signal: $\mathbf{h}_k^H \mathbf{w}_k \sqrt{\rho_k} s_k$
- Power: $\mathbb{E}[|s_k|^2] = 1$
- Result: $\rho_k |\mathbf{h}_k^H \mathbf{w}_k|^2$

DeepSeek-R1 Output

$$P_k^{\text{recv}} = p |\mathbf{h}_k^H \mathbf{w}_k|^2 \|\mathbf{w}_k\|^2$$

Errors:

- Confused pilot power p
- Redundant term $\|\mathbf{w}_k\|^2 = 1$
- Early mistake propagated

Error Example 4: Irrelevant System Mixing

Question: RIS-Assisted MIMO SINR

Derive the SINR for user k in RIS-assisted MIMO system:

$$\text{SINR}_k = [\text{MASK}]$$

Given: Single-user RIS-MIMO system, channel $\mathbf{h}_k$, transmit power P, noise σ^2

Correct Answer

$$\text{SINR}_k = \frac{P|\mathbf{h}_k^H \mathbf{w}_k|^2}{\sigma^2}$$

System:

- Single-user scenario
- No multi-user interference
- RIS for beamforming only

DeepSeek-R1 Output

$$\text{SINR}_k = \frac{P|\mathbf{h}_k^H \mathbf{w}_k|^2}{\sum_{j \neq k} P_j |\mathbf{h}_k^H \mathbf{w}_j|^2 + \sigma^2}$$

Errors:

- Injected NOMA interference $\sum_{j \neq k}$
- Assumed multi-user superposition (not in problem)

Key Takeaways: WirelessMathBench

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Main Findings

1. **Large Gap Exists:** Even best models achieve only 38% average accuracy, with FEC at 7.8%
2. **Reasoning Helps:** Explicit reasoning models (R1, o1) outperform general models by 5-8 points
3. **Difficulty Scaling:** Performance degrades exponentially with task complexity (roughly halves per level)
4. **Domain Training Insufficient:** Math specialization and telecom fine-tuning provide limited benefits
5. **Systematic Failures:** Models struggle with multi-variable coordination, symbol interpretation, and derivation structure

The Challenge

Can we build specialized models that perform better?

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Part 2

WirelessMathLM:

Teaching Mathematical Reasoning for LLMs in Wireless Communications with Reinforcement Learning

Xin Li, Mengbing Liu, Yiyang Zhu, Wenhe Zhang, Li Wei, Jiancheng An, Chau Yuen

Preprint

The Challenge: From Evaluation to Specialization

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Key Finding from Part 1

Even the **best model** (DeepSeek-R1) achieves only **38% average accuracy**
Full Equation Completion: **mere 7.8%** success rate

Our Goal

Train a **specialized model** for wireless mathematics

Challenge 1: Data

The Problem:

- Only 587 problems in Part 1
- Insufficient for robust training

Our Solution:

- **WirelessMathBench-XL**
- Scale up to 4,027 problems

Challenge 2: Training

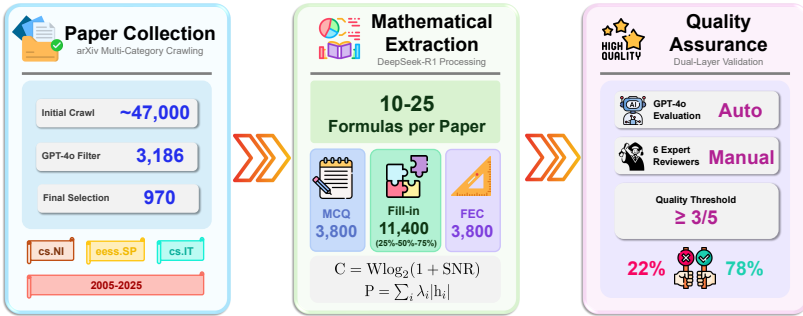
The Problem:

- SFT: Expensive expert annotations
- Large models: Huge resources

Our Solution:

- **Direct GRPO training**
- No SFT warm-start needed

Scaling Up: WirelessMathBench → XL



Dataset Expansion

Scale Increase:

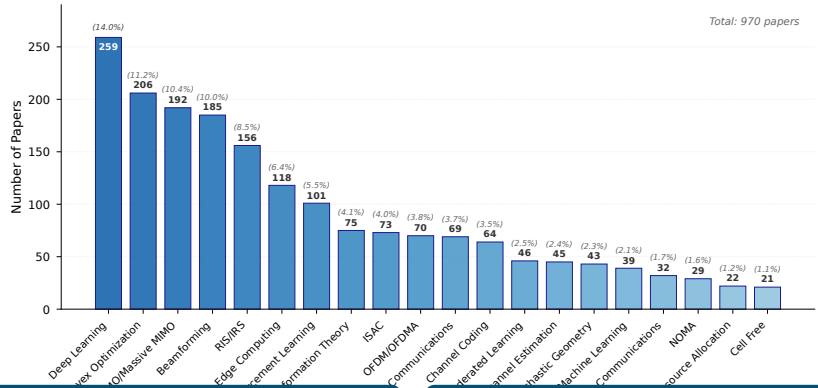
- 587 → **4,027 problems** (7×)
- 40 → **970 papers** (24×)

Quality Control

Dual-layer QA:

1. Automated GPT-4o scoring (1-5)
2. Expert validation ($\geq$ 2 required)

Technical Coverage: WirelessMathBench-XL



Top Techniques

- Convex Optimization (11.2%)
- MIMO/Massive MIMO (10.4%)
- RIS/IRS (8.5%)

Temporal Distribution

- 2005-2018 (3G/4G): 2.9%
- 2019-2023 (5G): 32.7%
- 2024-2025 (5G+): 64.4%

Why Reinforcement Learning for Math Reasoning?

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Limitations of Supervised Learning

Supervised Fine-Tuning (SFT):

- Requires complete solution traces
- Model mimics training data patterns
- Limited exploration of solution space
- Cannot go beyond training distribution

The Imitation Gap:

- Training: "Copy teacher's steps"
- Testing: Need creative problem-solving
- Result: Struggles with novel problems

RL Enables Exploration

Key Advantages:

- **Outcome-based learning**
  Focus on correctness, not steps
- **Solution space exploration**
  Discover multiple solving paths
- **Self-improvement**
  Learn from trial and error
- **Beyond human demonstrations**
  Can find novel solutions

Evidence from DeepSeek-R1

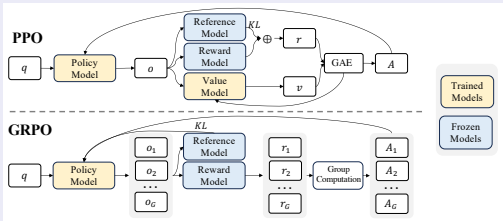
"RL transforms models from pattern matchers to problem solvers"

GRPO vs PPO: Architectural Comparison

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

How GRPO (Group Relative Policy Optimization) Simplifies RL

PPO: Complex Pipeline



Required Models:

- Reference Model (frozen)
- Policy Model (training)
- Value Model
- Reward Model

GRPO: Streamlined Design

Only 2 Components:

1. **Policy Model** (training)
2. **Reference Model** (reference)

Key Innovation: Group Comparison

- Sample G outputs per problem
- Compute group statistics:

$$A_i = \frac{r_i - \mu_G}{\sigma_G}$$

- No need for value model!
- Relative learning within group

Verification-Based Reward System

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Two-Component Reward Function

$$r(x, y) = \alpha \cdot r_{\text{format}}(y) + (1 - \alpha) \cdot r_{\text{accuracy}}(x, y)$$

where $\alpha = 0.1$ balances format compliance with correctness

Format Reward r_{format}

Ensures structural correctness:

$$r_{\text{format}}(y) = \mathbb{I}[\text{regex\_match}(y, \texttt{".*\\\\boxed\{.*\}.*"})]$$

Verification Steps:

1. Check proper LaTeX syntax
2. Verify `\boxed{}` final answer
3. Ensure parseable structure

Purpose:

- Enable automated evaluation
- Maintain output consistency
- Support downstream parsing

Verification-Based Reward System

Two-Component Reward Function

$$r(x, y) = \alpha \cdot r_{\text{format}}(y) + (1 - \alpha) \cdot r_{\text{accuracy}}(x, y)$$

where $\alpha = 0.1$ balances format compliance with correctness

Accuracy Reward r_{accuracy} — Multi-level Verification

Level 1: Direct Matching

- MCQ: Extract and compare letters
- Simple expressions: Exact string match

Level 2: Symbolic Verification

- Normalize expressions: Remove spaces, `\mathbf`, `\boldsymbol`
- Check mathematical equivalence
- Verify dimensional consistency

Level 3: Semantic Checking

- GPT-4.1-mini for complex expressions
- All-or-nothing evaluation

$$r_{\text{accuracy}} = \mathbb{I}[\text{correct answer}]$$

Implementation Details: Training Configuration

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Model Architecture

Base Models:

- Qwen2.5-series: 0.5B, 3B, 7B
- Direct training from base checkpoints

Training Data:

- 3,227 problems (80% split)

Computational Efficiency

- Hardware: 4× NVIDIA A6000 (48GB)
- Time: 14h (0.5B), 40h (3B), 61h (7B)

Optimization Setup

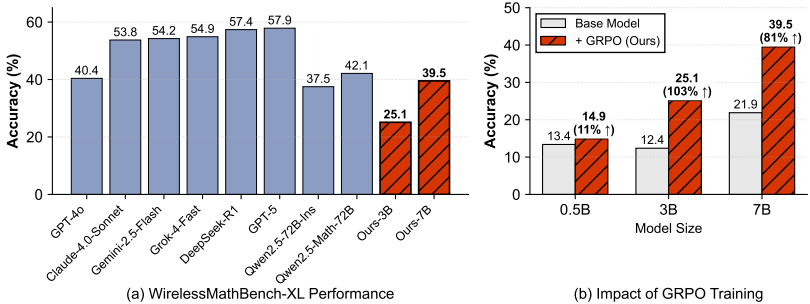
GRPO Parameters:

- Group size: $G = 8$
- Clip ratio: $\epsilon = 0.2$
- Temperature: $T = 0.6$ (validation)
- Temperature: $T = 1.0$ (training rollouts)

Training Schedule

- **Duration:** 40 epochs (240 steps)
- **Sequence length:** 2048 tokens max

Results: Dramatic Improvements



(a) WirelessMathBench-XL Performance

(b) Impact of GRPO Training

Key Achievements

- 7B model: 21.9% → 39.5% (**+17.6 pts, +81% relative**)
- 3B model: 12.4% → 25.1% (**+12.7 pts, +103% relative**)
- 0.5B model: 13.4% → 14.9% (+1.5 pts, +11% relative)

Detailed Results Breakdown

Model	MCQ	Fill-in	FEC	Overall	Gain
7B Models					
Qwen2.5-7B-Base	44.4%	14.3%	25.1%	21.9%	-
+ GRPO	53.4%	37.0%	36.1%	**39.5%**	+17.6
Relative Gain	*+20%*	*+159%*	*+44%*	*+81%*	
3B Models					
Qwen2.5-3B-Base	26.3%	7.1%	15.7%	12.4%	-
+ GRPO	48.9%	17.0%	28.8%	**25.1%**	+12.7
Relative Gain	*+86%*	*+139%*	*+83%*	*+103%*	
0.5B Models					
Qwen2.5-0.5B-Base	27.1%	5.3%	24.1%	13.4%	-
+ GRPO	30.1%	6.1%	26.2%	**14.9%**	+1.5
Relative Gain	*+11%*	*+15%*	*+9%*	*+11%*	

Surprising Discovery: Transfer to General Math

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Benchmark	Base	+GRPO	Gain	Type
7B Model Results				
MATH 500	52.0%	67.0%	+15.0	High School
Minerva-Math	12.1%	14.3%	+2.2	University
OlympiadBench	25.3%	30.2%	+4.9	Competition
AMC	27.7%	41.0%	+13.3	Competition
AIME24	6.7%	13.3%	+6.6	Competition
Average	**24.8%**	**33.2%**	**+8.4**	
3B Model Results				
MATH 500	41.6%	58.2%	+16.6	
Minerva-Math	5.9%	9.9%	+4.0	
OlympiadBench	14.7%	23.0%	+8.3	
AMC	18.1%	21.7%	+3.6	
Average	**16.0%**	**22.6%**	**+6.5**	

Understanding Positive Transfer

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Hypothesis 1: Skill Transfer

Wireless math requires:

- Matrix algebra
- Multi-step derivations
- Symbolic manipulation
- Constraint satisfaction

Evidence:

- Largest gains on MATH (+15pts)
- Strong on AMC (+13pts)
- Consistent across levels

Hypothesis 2: Deep Reasoning

GRPO training forces:

- Exploration of solution space
- Self-correction of errors
- Constraint verification
- Multi-step planning

Evidence:

- Improves on unseen problems
- Better at complex derivations
- Stronger error detection

Key Insight

Domain specialization done right can **strengthen fundamentals**

Qualitative Analysis: Solution Quality

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Analysis Scope

Comprehensive examination of 800 solutions generated by WirelessMathLM-7B on WirelessMathBench-XL test problems spanning all difficulty levels

What GRPO Training Achieved

1. **Structured Reasoning**
   - 99.1% systematic solutions
   - Clear logical connectives:"therefore" "thus" "hence"
2. **Knowledge Integration**
   - 87% correct problem identification
   - Physical + mathematical fusion
3. **Mathematical Sophistication**
   - Automatic constraint handling
   - Method justification
   - Physical intuition

Evidence of Real Understanding

- Not template filling
- Not simple pattern matching
- Genuine problem decomposition
- Context-aware reasoning
- Domain knowledge integration

Implication

Verification-based RL can develop sophisticated domain expertise without human feedback or supervised warm-start

Key Takeaways: WirelessMathLM

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Main Contributions

1. **Efficient Specialization:** 7B model approaches GPT-4o (39.5% vs 40.4%) using 100× fewer parameters than DeepSeek-R1
2. **Direct GRPO Works:** No supervised warm-start needed—verification-based RL sufficient for domain specialization
3. **Positive Transfer:** Deep domain training enhances general math (+8.4 pts average), contradicting catastrophic forgetting
4. **Consistent Scaling:** GRPO improves all model sizes (0.5B: +11%, 3B: +103%, 7B: +81%)
5. **Scalable Pipeline:** Semi-automated dataset construction from 47,000 papers

Broader Implication

Verifiable correctness enables efficient specialization in any technical domain:
Circuit design, Control theory, Cryptography, Formal verification, ...

Summary: Two Contributions

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

WirelessMathBench

ACL Findings 2025

The Problem:

- No benchmark for wireless math
- Unknown LLM capabilities

Our Solution:

- 587 expert-validated problems
- Progressive difficulty design
- 16 LLM evaluation

Key Finding:

- Best: 38% average, $<$8% FEC
- Clear need for specialization

`https://lixin.ai/WirelessMathBench`

WirelessMathLM

Preprint

The Challenge:

- How to train efficiently?
- Without massive resources?

Our Solution:

- 4,027 problems (970 papers)
- GRPO with verification
- Direct training from base

Key Achievements:

- 7B $\rightarrow$ 39.5% (near GPT-4o)
- $+$8.4pts general math

`https://lixin.ai/WirelessMathLM`

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Thank You!

WirelessMathBench
ACL Findings 2025

https://lixin.ai/WirelessMathBench

WirelessMathLM
Preprint

https://lixin.ai/WirelessMathLM

Questions & Discussion

Xin Li xin019@e.ntu.edu.sg https://lixin.ai

Open Questions & Discussion

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Seeking Your Insights

1. **Where should LLMs meet wireless communications?**
2. **What wireless problems are most "LLM-ready"?**
3. **How to leverage LLMs' language capabilities for wireless?**
4. **What does the community need most?**
   - Datasets? Tools? Pre-trained models? Benchmarks?
   - *How can we collaborate to accelerate progress?*

Prior Work in Wireless + LLMs

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Early Attempts (Focus on Protocols & Knowledge)

- **TeleQnA**: Q&A on 3GPP standards
- **WirelessLLM**: Knowledge retrieval
- **TelecomGPT**: Protocol understanding + basic formulas + codes
- **Tele-LLMs**: Fine-tuned on telecom corpus

What They Did Well

- Knowledge extraction
- Protocol summarization
- Basic code generation
- Standard definitions

What's Missing

- **No focus on mathematics**
- **No systematic evaluation**
- **No specialized math model**
- **No benchmark for reasoning**

→ We focus on mathematical reasoning, not just knowledge

Comparison with Related Work

Work	Domain	Approach	Data Size	Model Size
DeepSeekMath	General Math	SFT + RL	Large corpus	7B
Qwen2.5-Math	General Math	SFT + RL	Competitions	7B-72B
TelecomGPT	Wireless	SFT	Protocols	8B
LLaMA-3-Tele	Wireless	SFT	Telecom corpus	8B
WirelessMathLM	**Wireless Math**	**Direct GRPO**	**4,027 problems**	**0.5B-7B**

Our Unique Features

- No SFT warm-start needed
- Verification-based rewards
- Research-level problems
- Transfer to general math

Advantages

- More efficient training
- Domain-specific focus
- Verifiable correctness
- Broader applicability