

ABSTRACT

**for the dissertation of Yerlan Makhmut
submitted for the degree of Doctor of Philosophy (PhD) in the
educational program 8D06105 – System Engineering
on the topic: “Development of hybrid high-performance algorithms for numerical
simulation of oil displacement problems”**

Relevance of the research topic. Accurate modelling of multiphase flow with chemical transport in porous media is critically important for engineering applications such as petroleum reservoir simulation, enhanced oil recovery, and the sustainable management of mature oil fields. These processes are governed by a key class of differential equations expressing the laws of conservation of mass and momentum, and their faithful numerical reproduction is essential for the planning of injection scenarios, well placement, and economic evaluation of recovery projects.

A representative example of such flow is polymer flooding, in which a polymer-laden aqueous phase is injected into a hydrocarbon reservoir to improve the mobility ratio and to delay the breakthrough of the displacing front. The governing equations of two-phase flow with polymer transport form a system of three strongly coupled nonlinear partial differential equations — for the oil phase, the water phase, and the dissolved polymer — supplemented by Darcy’s law, the Brooks–Corey relative permeability model, a power-law polymer-viscosity modification, and a Langmuir adsorption isotherm. The strong nonlinearity, the sharp displacement fronts, and the adsorption-driven nonlinear coupling between the dissolved and adsorbed polymer concentrations pose significant numerical challenges for traditional sequential solvers.

Numerical methods such as finite-difference and finite-volume IMPES schemes are widely used in reservoir simulation; however, when applied to high-resolution polymer-flooding problems at field scale, they require very fine spatial and temporal discretisation, leading to prohibitive computational costs. These limitations have stimulated the search for more efficient computational approaches, and in particular for parallel algorithms capable of exploiting the full hierarchy of contemporary heterogeneous computing hardware — distributed-memory clusters, shared-memory multi-core processors, and GPU accelerators.

High-performance computing methods address this problem by mapping the governing physics solvers onto modern parallel hardware while preserving numerical correctness. In recent years, hybrid combinations of MPI, OpenMP, and CUDA have demonstrated significant speedup potential for coupled multiphase-flow problems. However, the majority of existing studies are limited to single-paradigm parallelism or to a single GPU; within the polymer-flooding framework

specifically, a significant gap remains in the simultaneous optimisation of all three coupled physics kernels through on-chip memory tiling, and in the coherent integration of three independent levels of parallelism within a single computational pipeline. The present dissertation proposes a unified hierarchy of seven parallel algorithms capable of solving the coupled polymer-flooding system across the full range of contemporary heterogeneous hardware, offering new momentum for the high-performance modelling of enhanced oil recovery processes.

Objective of the dissertation

The purpose of this work is to develop hybrid computational algorithms for solving oil displacement problems in oil reservoirs by using high-performance computing technologies, and to compare their effectiveness systematically in terms of speedup, scalability, and numerical accuracy against a validated sequential baseline.

Research tasks:

- to develop a serial numerical algorithm for the one-dimensional polymer-flooding problem and to establish it, after validation, as the performance baseline;
- to implement and evaluate CPU-based parallel algorithms — OpenMP shared-memory, MPI distributed-memory, and hybrid MPI+OpenMP — on multi-core systems across six problem sizes;
- to implement a CUDA global-memory GPU algorithm, advance it to a novel CUDA shared-memory GPU pipeline, and extend the implementation to a multi-GPU MPI+CUDA configuration;
- to develop a triple-hybrid MPI+OpenMP+CUDA algorithm and to benchmark every implementation against the serial baseline using Intel VTune, mpiP, NVIDIA Nsight Compute, and NVIDIA Nsight Systems.

Research methods: the research methods include literature review, mathematical modelling, finite-difference discretisation, the IMPES time-stepping scheme with adaptive CFL control, the design and implementation of parallel algorithms using the OpenMP, MPI, and CUDA programming models, numerical experiments across six grid sizes ranging from $N = 65,536$ to $N = 67,108,864$ cells, as well as accuracy evaluation and validation of the results through comparison with the analytical Buckley–Leverett solution, mass-conservation tests against machine precision, and profiling-tool corroboration using Intel VTune, mpiP, NVIDIA Nsight Compute, and NVIDIA Nsight Systems.

Object of the research

Two-phase (water–oil) flow with dissolved-polymer transport in one-dimensional porous media and the system of three coupled partial differential equations describing it.

Subject of the research

Methods for the high-performance parallel solution of the coupled pressure, saturation, and polymer-concentration fields in polymer flooding using OpenMP, MPI, CUDA, and hybrid combinations thereof on heterogeneous CPU–GPU computing platforms.

Scientific novelty of the results:

- A fully coupled GPU-based numerical algorithm for pressure, saturation, and polymer concentration was developed using a shared-memory tiling strategy, in which warp-level shuffle reductions eliminate inter-thread synchronisation overhead and SM-resident tiles amortise global-memory latency, collectively improving computational efficiency in multiphase porous-media simulations.
- A scalable triple-hybrid MPI–OpenMP–CUDA parallel algorithm was developed for large-scale oil displacement simulations, enabling efficient distributed multi-GPU computations on significantly larger computational grids.

Provisions to be defended:

- A fully coupled GPU implementation of the pressure, saturation, and polymer-concentration solvers based on shared-memory tiling and warp-level parallelism achieved up to $20.2\times$ speedup, while approaching near-peak FP64 computational performance.
- The developed triple-hybrid MPI+OpenMP+CUDA parallel algorithm demonstrated high scalability and achieved up to $34.7\times$ speedup compared with the optimised serial implementation for large-scale oil displacement simulations.

Theoretical and practical significance of the research

Theoretical significance: the research contributes to the development of the theoretical foundations of parallel algorithm design for coupled multiphase-flow problems and extends the methodological capabilities of full-pipeline shared-memory tiling and three-level hybrid parallelism in the simulation of porous-media transport.

Practical significance: the developed algorithms reduce the computational cost of high-resolution polymer-flooding simulations and improve the efficiency of reservoir simulation software on widely accessible workstation-class hardware, supporting timely engineering decisions in enhanced oil recovery.

Reliability of scientific results

The reliability of the results is ensured by comparison with the analytical Buckley–Leverett solution, validation against the serial baseline to a relative error below 10^{-6} in the pressure, saturation, and polymer-concentration fields, verification of mass conservation to machine precision, consistency with the underlying physical laws, and independent corroboration of the performance measurements through the Intel VTune, mpiP, NVIDIA Nsight Compute, and NVIDIA Nsight Systems profiling tools.

Relevance to research programs

This dissertation was carried out within the framework of the research project “High-performance computer modelling of multiphase fluid flow in porous media under uncertainty” (State registration No. AP09260564, grant funding, Al-Farabi Kazakh National University, 2021–2023).

Volume and structure of the dissertation

The dissertation consists of an introduction, eight main chapters, a conclusion, and a list of references.

The first chapter presents the research motivation, problem statement, aim and research tasks, scientific novelty, statements submitted for defense, scope and limitations, and the experimental platform. The second chapter reviews the relevant literature in four domains — polymer flooding as an enhanced oil recovery technology, mathematical modelling of multiphase flow in porous media, numerical methods for reservoir simulation, and high-performance computing approaches to reservoir simulation — and identifies the three literature gaps closed by the dissertation. The third chapter presents the mathematical model: the three governing partial differential equations, the constitutive relations, and the one-dimensional boundary-condition formulations.

The fourth chapter develops the finite-difference discretisation, the IMPES time-stepping scheme with adaptive CFL control, and the four sub-algorithms of the sequential simulator. The fifth chapter presents the three CPU-based parallel algorithms (OpenMP, MPI, and hybrid MPI+OpenMP) with full performance evaluation across all six grid sizes. The sixth chapter develops the CUDA global-memory and CUDA shared-memory algorithms, including the first scientific novelty of the dissertation. The seventh chapter presents the two multi-GPU hybrids (MPI+CUDA and the triple-hybrid MPI+OpenMP+CUDA), including the second scientific novelty. The eighth chapter reports the unified seven-algorithm benchmarking and validation, with profiling-tool corroboration.

Publications based on the research results

A total of 10 publications were produced on the topic of the dissertation, including 4 in journals recommended by the Science and Higher Education Quality Assurance Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan, 4 in peer-reviewed journals indexed in Web of Science and/or Scopus, and 2 in international conference proceedings.

Scopus:

1. Yerlan Makhmut, Timur Imankulov, Sergei Gorlatch, Bazargul Matkerim. A CUDA Performance Study of Global- and Shared-Memory Kernels for the Buckley–Leverett Polymer-Flooding Problem // Applied Sciences. – May 2026. – Q1 (79%).
2. Imankulov T., Kenzhebek Y., Bekele S.D., Makhmut E. Enhancing Oil Recovery Predictions by Leveraging Polymer Flooding Simulations and Machine Learning Models on a Large-Scale Synthetic Dataset // Energies. – 2024. – Vol. 17, No. 14, 3397. – Q1 (90%).
3. Daribayev B., Mukhanbet A., Makhmut E., Azatbekuly N., Imankulov T. Quantum–Classical Hybrid Architecture for Solving the Muskat–Leverett Model // IET Quantum Communication. – 2025. – Vol. 6, No. 1, e70022. – Q1 (79%).
4. Makhmut E., Imankulov T., Daribayev B. Parallel Implementation of Muskat–Leverett Equation Using CUDA // Journal of Mathematics, Mechanics and Computer Science. – 2023. – Vol. 119, No. 3. – Q4 (18%).

International Conferences:

1. Makhmut E., Imankulov T., Daribayev B., Akhmed-Zaki D. Development of Hybrid Parallel Computing Models to Solve Polymer Flooding Problem // European Association of Geoscientists & Engineers, ECMOR 2022. – September 2022. – Vol. 2022. – P. 1–14. (Scopus).
2. Imankulov T., Kenzhebek Y., Makhmut E., Akhmed-Zaki D. Using Machine Learning Algorithms to Solve Polymer Flooding Problem // European Conference on the Mathematics of Geological Reservoirs (ECMOR 2022). – 2022. (Scopus).

Science and Higher Education Quality Assurance Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan:

1. Makhmut E., Kenzhebek Y.G., Imankulov T.S. Solving the Oil Displacement Problem Using a Hybrid MPI+OpenMP Parallel Algorithm // Bulletin of Abai KazNPU, Series of Physical and Mathematical Sciences. – 2023. – No. 1(81). – P. 56–65.
2. Makhmut E., Imankulov T.S., Matkerim B. Design and Development of a Hybrid (MPI+CUDA) Parallel Program for Solving the Oil Displacement Problem // Bulletin of the National Engineering Academy of the Republic of Kazakhstan. – 2024. – No. 1(91).

3. Makhmut E., Imankulov T. Comparative Analysis of Reduction Algorithms in CUDA for the Pressure Equation // Journal of Problems in Computer Science and Information Technologies. – 2023. – Vol. 1, No. 2.

4. Mustafin M.B., Makhmut E. Comparative Analysis of the Efficiency of Graphics Processing Units in Machine Learning // Bulletin of the National Engineering Academy of the Republic of Kazakhstan. – 2024. – December.