Accelerating Kinetic Simulations of Electrostatic Plasmas with Reduced-Order Modeling

Ping-Hsuan Tsai  
Department of Computer Science  
University of Illinois Urbana–Champaign  
Urbana, IL 61801  
pht2@illinois.edu

Seung Whan Chung  
Lawrence Livermore National Laboratory  
Livermore, CA 94550  
chung28@llnl.gov

Debojyoti Ghosh  
Lawrence Livermore National Laboratory  
Livermore, CA 94550  
ghosh5@llnl.gov

John Loffeld  
Lawrence Livermore National Laboratory  
Livermore, CA 94550  
loffeld1@llnl.gov

Youngsoo Choi  
Lawrence Livermore National Laboratory  
Livermore, CA 94550  
choi15@llnl.gov

Jonathan L. Belof  
Lawrence Livermore National Laboratory  
Livermore, CA 94550  
belof1@llnl.gov

Abstract

Despite the advancements in high-performance computing and modern numerical algorithms, the cost remains prohibitive for multi-query kinetic plasma simulations. In this work, we develop data-driven reduced-order models (ROM) for collisionless electrostatic plasma dynamics, based on the kinetic Vlasov-Poisson equation. Our ROM approach projects the equation onto a linear subspace defined by principal proper orthogonal decomposition (POD) modes. We introduce an efficient tensorial method to update the nonlinear term using a precomputed third-order tensor. We capture multiscale behavior with a minimal number of POD modes by decomposing the solution into multiple time windows using a physical-time indicator and creating a temporally-local ROM. Applied to 1D–1V simulations, specifically the benchmark two-stream instability case, our time-windowed reduced-order model (TW–ROM) with the tensorial approach solves the equation approximately 280 times faster than Eulerian simulations while maintaining a maximum relative error of 4% for the training data and 13% for the testing data.

1 Introduction

Kinetic modeling of collisionless electrostatic plasma relies on the Vlasov–Poisson equations that describe charged particle distribution under self-consistent electrostatic fields. Due to high dimensionality, scale disparities, and nonlinearity, solving these equations is computationally challenging. Lagrangian particle-in-cell (PIC) methods [1] are often used, where particles are advanced along the characteristic curves of the Vlasov equation [2]. An alternative is the grid–based Eulerian methods that allow the use of advanced numerical algorithms for partial differential equations (PDEs) [3–5]. In this paper, we consider the latter as full-order models (FOMs). While high-performance computing and advanced algorithms enable high-fidelity kinetic plasma simulations, they remain intractable for parametric studies. In this work, we propose to address this by data-driven projection-based reduced-order models, which have been successfully applied to many physical simulations, e.g., fluid...
dynamics [6–13], nonlinear diffusion [14], Boltzmann transport [15], design optimization [16–19].
Starting with high-resolution Eulerian simulations, we derive a low-dimensional surrogate model
that is solved at a reduced computational cost while providing approximate solutions with acceptable
accuracy.
We consider the 1D–1V parametric Vlasov–Poisson equation discretized with a high-order conser-

vative finite difference method. Traditional reduced-order modeling faces challenges of stability,
efficiency, and accuracy due to the problem’s multi-scale nature, limiting its benefits. We overcome
this by partitioning the solution into multiple time windows and creating temporally-local ROMs [20–
24]. The primary computational expense arises from the nonlinear hyperbolic term evaluation in the
high-dimensional space. To improve efficiency, we introduce a tensorial approach for updating this
term using a precomputed third-order tensor. Our numerical experiments demonstrate the efficacy
of the proposed time-windowed ROM (TW–ROM) with the tensorial approach, and we achieve
significant speed improvements.

2 Reduced-order modeling of the 1D–1V Vlasov–Poisson equations

In Section (2.1), we introduce the parametric 1D–1V Vlasov–Poisson equation and the initial and
boundary conditions that give rise to the two-stream instability. In Section (2.2), we briefly introduce
the full-order model. In Section (2.3), we present the framework for the time-windowing reduced-
order model and the tensorial approach. The reduced order model is constructed in the offline phase
deployed in the online phase.

2.1 Parametric 1D–1V Vlasov–Poisson equations

We consider the parametric 1D–1V Vlasov–Poisson equations,

\begin{align}
\partial_t f(x, v, t) + v \partial_x f(x, v, t) + E(x, t) \partial_v f(x, v, t) &= 0, \\
\partial_t^2 \phi(x, t) &= \int f(x, v, t) \, dv - \int f(x, v, t) \, dv \, dx, \\
f(x, v, 0) &= \frac{8}{\sqrt{2\pi t}} \left(1 + \alpha \cos \left(2k\pi \frac{x}{L}\right)\right) \left[\exp \left(-\frac{(v-v_0)^2}{2T}\right) + \exp \left(-\frac{(v+v_0)^2}{2T}\right)\right].
\end{align}

Here \( f(x, v, t) \) is the plasma distribution function, \( \phi(x, t) \) is the electrostatic potential, and
\( E(x, t) = -\partial_v \phi(x, t) \) is the electric field. The spatial and velocity coordinates are \( x \) and \( v \), respectively.
The simulation time interval is \( [0, t_f] \) where \( t_f \in \mathbb{R}_+ \) is the final time, and the phase-space domain
is \( (x, v) \in \Omega := [0, 2\pi] \times [-3.5v_0, 3.5v_0] \) with periodic boundaries in \( x \) and homogeneous Dirichlet
boundaries in \( v \). The initial distribution \( f(x, v, 0) \) is characterized by three parameters: the tempera-
ture \( T \), the perturbation amplitude \( \alpha \), and the initial stream velocity \( v_0 \) with \( k = 1 \) and \( L = 2\pi \). We
focus on studying the effect of parameterized initial distributions on the plasma dynamics.

2.2 Full-order model (FOM)

The semi-discretization of Eq. (1) in space can be written as:

\[ \frac{df(t; \mu)}{dt} = G(f, t; \mu), \quad t \in [0, t_f], \quad \text{with} \quad f(0; \mu) = f_0(\mu), \tag{2} \]

where \( \mu = (T, \alpha, v_0) \) is the parameter in a parameter domain \( \mathcal{D} \); \( f \in \mathbb{R}^{N_f} \) denotes the parameterized
time-dependent solution to the dynamical system with an initial state \( f_0 \) and \( N_f \) is the FOM degrees
of freedoms; \( G \) represents the nonlinear matrix function coming from the hyperbolic term in (1).

We discretize Eq. (1) in space with a high-order conservative finite-difference code on a Cartesian grid.
The domain \( \Omega \) is discretized on a \( 256 \times 256 \) grid leads to \( N_f = 65,536 \). The spatial derivatives
are computed using the fifth-order weighted essentially nonoscillatory (WENO) scheme [25]. The
electrostatic potential is computed by solving the periodic Poisson equation with the fast Fourier
transformation (FFT). The classical four-stage fourth-order explicit Runge–Kutta time integrator is
used to evolve the resulting ordinary differential equation (ODE), given by Eq. (2) with a uniform time
step of \( \Delta t = 0.0025 \).
2.3 Time-windowing reduced-order model (TW-ROM)

The formation of the two-stream instability occurs due to the nonlinear evolution of (1). This instability is a well-known phenomenon in plasma physics, which is generated by two counterstreaming beams. In this process, the kinetic energy of particles excites a plasma wave, which then transfers to electrostatic potential energy [26]. The main interest of numerical studies in the two-stream instability is to investigate how the parameter $\mu$ affects the solutions and the growth rate of instability. The primary difficulty for the ROM are the three stages of the solutions, i.e., short transient, growth, and statistically stationary stages, which require many reduced bases for the ROM to accurately capture the behavior in each stage [27].

We propose a framework to overcome these difficulties by employing multiple reduced models in time. The idea of the methodology is to construct local ROMs in the parameter-time domain using physical time $t$ as the indicator for clustering and classification [20,23]. Notice that the methodology is not limited to the 1D-1V Vlasov-Poisson equations and has been applied to Euler equations [20] and Navier-Stokes equations [28]. For a generic problem parameter $\mu \in \mathcal{D}$, let $t_f$ be the final time in the ROM simulation. The computation in the online phase is performed using different reduced bases in $N_w$ subintervals of the temporal domain $[0, t_f]$, i.e., $0 = t_0 < t_1 < \cdots < t_{N_w-1} < t_{N_w} = t_f$. With the partition of the indicator range, instead of directly assembling all the snapshot samples into a single huge snapshot matrix, the FOM states are first classified into groups. Let $m \in \mathbb{N}(N_w)$ be a group of index. We denote the subset of the index of time whose corresponding snapshot belongs to the $m$-th group as

$$
\mathcal{G}_m = \{n \in \mathbb{Z} : 0 \leq n \leq N_t \text{ and } t_n \in [t_{m-1}, t_m)\},
$$

where $N_t$ is the number of time steps and $t_{N_t} = t_f$. Then the snapshot matrix of the distribution $U_{f,m}$ in the $m$-th group is formed by assembling the corresponding snapshots, i.e., $U_{f,m} \equiv [\phi_n(\mu)]_{n \in \mathcal{G}_m}$. Therefore, for $t \in T_j \equiv [t_{j-1}, t_j]$, we employ the reduced bases constructed from the snapshot group $\mathcal{G}_m$. More precisely, we use the solution representation

$$
\tilde{f}(t; \mu) = \Phi_{f,m}^T \tilde{f}_j(t; \mu), \quad \text{where } \Phi_{f,m}^T \equiv \left[\phi_{1,f}^m \cdots \phi_{n_f,f}^m\right] \in \mathbb{R}^{N_j \times n_f}.
$$

Here $\Phi_{f,m}^T$ is the distribution solution basis matrix constructed using Proper orthogonal decomposition (POD) [29] from $U_{f,m}$ with $\phi_{i,f}^m$ being the $i$-th reduced basis vector and $\tilde{f}_j : T_j \times \mathcal{D} \rightarrow \mathbb{R}^{n_f}$ are the time-dependent generalized coordinates for distribution field in the time interval $T_j$ with $n_f$ being number of distribution reduced basis vectors. The reduced order model for each time interval $T_j$ is derived by replacing $f$ with $\tilde{f}$ in (2) and employ Galerkin projection to close the system:

$$
\frac{d\tilde{f}_j}{dt} = (\Phi_{f,m}^T)^T \mathbf{G}(\Phi_{f,m}^T \tilde{f}_j, t; \mu).
$$

Here, we use the assumption of basis orthonormality. The initial condition is given by projecting onto the ROM spaces, i.e.,

$$
\tilde{f}_j(t_{j-1}; \mu) = (\Phi_{f,m}^T)^T \hat{f}(t_{j-1}; \mu).
$$

2.3.1 Tensorial approach

There is one major issue with (3). The nonlinear matrix function, $\mathbf{G}$, changes every time the state variables evolve. Additionally, it needs to be multiplied by the basis matrix whenever the updates in the nonlinear term occur, which scales with the FOM size $N_f$. Therefore, we cannot expect any speed-up without special treatment of the nonlinear term. To overcome this issue, we introduce a tensorial approach to efficiently update the nonlinear term using a precomputed third-order tensor. This requires solving an additional reduced system for the Poisson equation in time interval $T_j$:

$$
-(\Phi_{\phi}^m)^T \nabla^2 (\Phi_{\phi}^m \phi^j) = (\Phi_{\phi}^m)^T \int (\Phi_{\phi}^m \tilde{f}_j) dv - (\Phi_{\phi}^m)^T \int (\Phi_{\phi}^m \tilde{f}_j) dv dx,
$$

where $\Phi_{\phi}^m$ is the potential solution basis matrix constructed using POD from $U_{\phi,m} \equiv [\phi_n(\mu_k)]_{n \in \mathcal{G}_m}$ and $\phi^j : T_j \times \mathcal{D} \rightarrow \mathbb{R}^{n_\phi}$ are the time-dependent generalized coordinates for potential field with $n_\phi$ being the number of potential reduced basis vectors. With the relation $E = -\nabla \phi$, an approximated electric field $\tilde{E}^j$ can approximated:

$$
\tilde{E}^j(x,t) = -\nabla (\Phi_{\phi}^m \phi^j) = \sum_{i}^{n_f} \hat{\phi}_i^j \nabla \phi_{i_E,t}^m = \sum_{i}^{n_f} \hat{\phi}_i^j \phi_{E,t}^m.
$$
The relative errors at the four training data (displayed on four corners) are less than 13%. The maximum relative error for testing data is 63%. To evaluate the TW-ROM performance, the relative error of the approximated solution is measured against the corresponding FOM solution at the final time $f$. The TW-ROM is able to capture the growth rate of the interpolation cases and extrapolation cases ($\mu = (0.08, 0.001, 1)$ and $(0.07, 0.0025, 1)$) and the extrapolation cases ($\mu = (0.08, 0.0015, 1)$ and $(0.07, 0.0025, 1)$) as well.

Fig. 1a further displays the predicted distribution $f$ of the TW-ROM at time instance $t = 8$ and $t = 10$ for three $\mu$ values; $\mu = (0.095, 0.00225, 1)$ and $\mu = (0.08, 0.0015, 1)$ as interpolation cases, and $\mu = (0.07, 0.0025, 1)$ as an extrapolation case. Despite some discrepancies in the predicted $f$.
values for $\mu = (0.08, 0.0015, 1)$ and $\mu = (0.07, 0.0025, 1)$, the TW-ROM accurately capture the occurrence of instability.

For a given parameter $\mu$, the wall-clock time for the FOM is about 56 (seconds), whereas the wall-clock time for the TW-ROM is about 0.2 (seconds), achieving a 280 times speed-up and a 4,480 times speed-up in CPU hours. The construction of the tensor $G_2$ for 100 time windows takes about 87.2 (seconds) which is longer than one FOM simulation. However, the speed-up gained in the online stage compensates the overhead. Furthermore, the current implementation of the tensor construction is serial, suggesting the potential for further reduction in construction time through parallel implementation.

The testing and training data (FOM solutions) and the TW-ROM offline and online phases are performed on Livermore Computing Quartz [30] with Intel Xeon CPUs, 128 GB memory, peak TFLOPS of 3251.4, and peak single CPU memory bandwidth of 77 GB/s. We use the functionalities in libROM [31], an existing asset with Apache license, to construct the TW-ROM.

A potential constraint lies in the choice of the indicator for partitioning the temporal domain. The existing TW-ROM is established using a physical-time indicator, resulting in a temporal partition that is independent of parameters. Figs. 1a-1b demonstrate that the TW-ROM can effectively approximate solutions and capture growth rates within a reasonable range of $T$ and $\alpha$. Moreover, it can readily adapt to other parameterization types. However, to extend the application of TW-ROM to a broader or higher-dimensional parameter space, a more intelligent indicator based on physics is essential.

In the context of 2D-2V and 3D-3V problems, the solution dynamics are anticipated to become even more intricate. This complexity requires standard ROM to use large number of modes to accurately represent the dynamic behavior. Consequently, the development of a time-windowing ROM becomes essential. This approach is crucial not only for capturing the multi-scale behavior but also for maintaining computational efficiency.
4 Broader impact

This paper presents a novel data-driven reduced-order modeling approach employing time-windowing and a tensorial strategy to expedite kinetic simulations of electrostatic plasmas. The envisioned framework is anticipated to exert a significant influence on the computational science community and holds the potential for diverse applications across various engineering and scientific domains. It is important to note that this work has been conducted with careful consideration, and no adverse ethical or societal consequences are associated with its findings.

Acknowledgments and Disclosure of Funding

This work was performed under the auspices of the U.S. Department of Energy (DOE), by Lawrence Livermore National Laboratory (LLNL) under Contract No. DE-AC52–07NA27344 and was supported by Laboratory Directed Research and Development funding under projects 21-SI-006 and 22-SI-006. Y. Choi was also supported for this work by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, as part of the CHaRMNET Mathematical Multifaceted Integrated Capability Center (MMICC) program, under Award Number DE-SC0023164. IM release: LLNL-CONF-855151.

References


