Bachelorarbeit
zur Erlangung des akademischen Grades
Bachelor of Science

Zachary Schellin
376930

December 30, 2020

Contents

1 Introduction

The Bhatnagar, Gross, Krook equation (BGK) is a kinetic collision model of ionized and neutral gases valid for rarefied as well as other pressure regimes [1]. Generating data of such a flow field is essential for various industry and scientific applications[**REF**]. With the intention to reduce time and cost during the data generating process, experiments were substituted with computational fluid dynamics (CFD) computations. Consequently reduced-order models (ROMs) coupled to aforementioned computations were introduced to further the reduction of time and cost. The thriving field of artificial intelligence operates in model order reduction for data visualization/analysis since the 80's (Quelle?) and has now surfaced in fluid mechanics. This thesis will cover the use of artificial intelligence for model order reduction in fluid mechanics.

1.1 State of the art

State of the art model reduction of dynamical systems is done via proper orthogonal decomposition (POD) which is an algorithm feeding on the idea of singular value decomposition (SVD)[2][3]. POD captures a low-rank representation on a linear manifold. So called POD modes, derived from SVD, describe the principle components of a problem which can be coupled within a Galerkin framework to produce an approximation of lower rank.

$$f(x) \approx \tilde{f}(x) \qquad \text{with} \qquad rk(f(x)) \gg rk(\tilde{f}(x)) \tag{1}$$

Bernard et al. use POD-Galerkin with an additional population of their snapshot database via optimal transport for the proposed BGK equation, bisecting computational run time (cost) in conjunction with an approximation error of 1 % [4]. Artificial intelligence in the form of autoencoders replacing the POD within a Galerkin framework is evaluated against the POD performance by Kookjin et al. for advection-dominated problems[5] resulting in sub 0.1% errors. An additional time inter- and extrapolation is evaluated. Using machine learning/ deep learning for reduced order modeling in CFD is a novel approach although "the idea of autoencoders has been part of the historical landscape of neural networks for decades"[6, p.493]. Autoencoders, or more precisely learning internal representations by the delta rule (backpropagation) and the use of hidden units in a feed forward neural network architecture, premiered by Rumelhart et al. (1986) [7]. Through so called hierarchical training Ballard et al.(1987) introduce a strategy to train auto autoassociative networks (nowadays referred to as autoencoders), in a reasonable time promoting further development despite computational limitations [8]. The so called bottleneck of autoencoders yields a non-smooth and entangled representation thus beeing uninterpretable by practitioners[9] leading to developements in this field. Rifai et al. introduce the contractive autoencoder (CAE) for classification tasks (2011), with the aim to extract robust features which are insensitive to input variations orthogonal to the low-dimensional non-linear manifold by adding a penalty on the frobenius norm of the instrinsic variables with respect to the input, surpassing other classification algorithms [9]. Subsequent development emerges with the manifold tangent classifier (MTC) [10]. A local chart for each datapoint is obtained hence characterizing the manifold which in turn improves classification performance. On that basis a generative process for the CAE is developed. Through movements along the manifold with directions defined by the Ja-

cobian of the bottleneck layer with respect to the input $\vec{x}_m = JJ^T$, sampling is realized [11].

1.2 Objective of this thesis

Due to the non-linearity of transport problems in particular shock fronts, the construction of a robust ROM for those cases poses several challenges. Proper orthogonal decomposition (POD) and it's numerous variants like shifted-POD[**?**], POD-Galerkin[**?**], POD+I [**?**] to name only a few of them, try to solve this problem by.....

1.3 Thesis outline

2 The BGK Equation

The Knudsen number eq. (2) introduced by Danish physicist Martin Knudsen is a measure for the rarefaction of gases. In eq. (2) λ represents the mean free path and L the characteristic length [4] of a particle. The mean free path describes the average distance a particle may travel between successive impacts [**WIKI**]. For idealized gases the mean free path can be calculated via eq. (3). In eq. (3) k_b is the Boltzman constant, p is the total pressure, T is the thermodynamic temperature and d is the hard shell diameter [**WIKI**].

$$Kn = \frac{\lambda}{L} \qquad (2) \qquad \lambda = \frac{k_b T}{\sqrt{2}\pi d^2 p} \qquad (3)$$

For Kndusen numbers $Kn > 10^{-2}$ collisions are predominant in comparison to free transport, whereas for $Kn < 10^{-2}$ free transport becomes the predominant behavior compared to collision[4]. This difference in turn characterizes flows where the Bolzmann equation (collisions) or the Navier-Stokes equations (free transport) are valid. Hence the former eq. (4) describes the dynamics of a gas flow, where f is the probability density distribution function for a particle at point $\mathbf{x} \in \mathbb{R}^3$ with velocity $\xi \in \mathbb{R}^3$ at time $t \in \mathbb{R}$.

$$\partial_t f(\mathbf{x}, \xi, t) + \xi \Delta_x f(\mathbf{x}, \xi, t) = Q(f, f) \qquad (4)$$

Originally Q is often the binary Boltzmann collision term, which can be intractable in pratice [1]. Thus the BGK equation utilizes the BGK-Operator for the collision term Q(f,f) eq. (5)[4].

$$Q(f, f) = \frac{M_f(\mathbf{x}, \xi, t) - f(\mathbf{x}, \xi, t)}{\tau(\mathbf{x}, t)} \qquad (5)$$

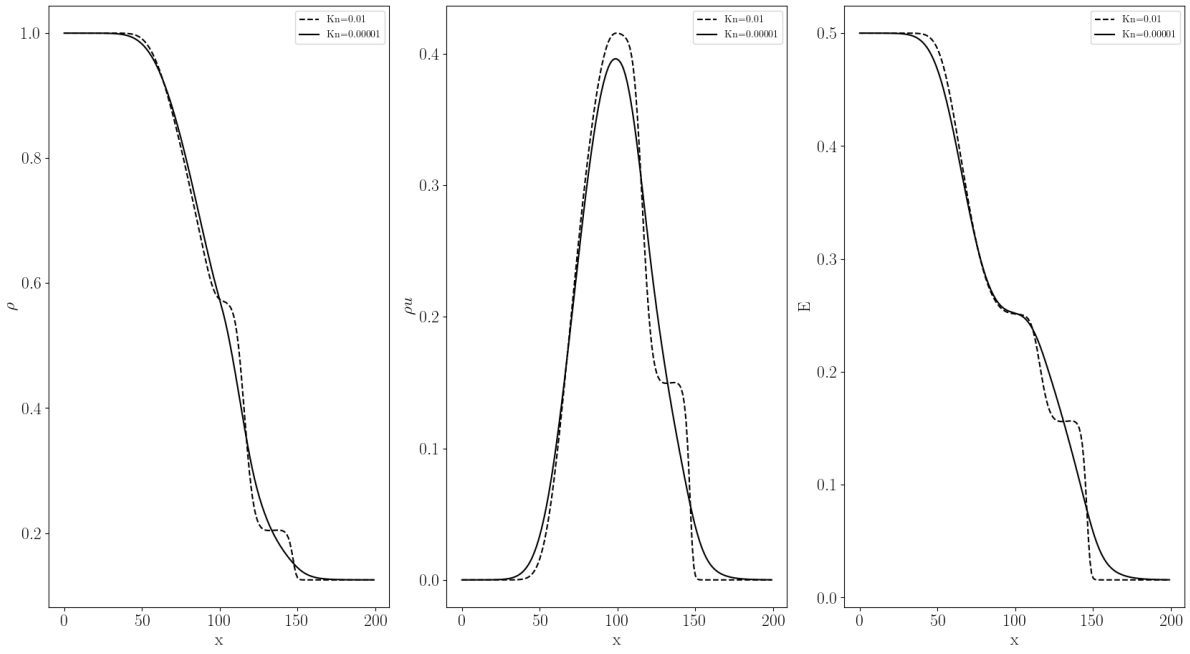
$$M_f(\mathbf{x}, \xi, t) = \frac{\rho(\mathbf{x}, t)}{(2\pi T(\mathbf{x}, t))^{\frac{3}{2}}} \exp(-\frac{||\xi - U(\mathbf{x}, t)||^2}{2T(\mathbf{x}, t)}) \qquad (6)$$

$$\tau(\mathbf{x},\mathbf{t}) = \frac{Kn}{\rho(\mathbf{x}, t) T^{1-\nu}(\mathbf{x}, t)} \qquad (7)$$

For many kinetic gas problems it is sufficient to replace the complexity of the Boltzmann collision term by a mean-free-path approach. τ is referred to as the relaxation time, considering the fact, that collisions tend to relax the distribution function to an equilibrium state f_0, where $f_0 = M_f$ the equilibrium state is approximated by a Maxwellian distribution function[1]. In eq. (6) $U(\mathbf{x}, t) = (u(\mathbf{x}, t), v(\mathbf{x}, t), w(\mathbf{x}, t))^T$ is the macroscopic velovity, $T(\mathbf{x}, t) \in \mathbb{R}$ and $\rho(\mathbf{x}, t) \in \mathbb{R}$is the temperature and density of the gas respectively. In eq. (7) $\nu \in \mathbb{R}$ is the exponent of the viscosity law of the gas. As a result one obtains th BGK equation which can be utilized for both hydrodynamic and rarefied regimes and meets global conservation as discussed in section 2.1. By multiplying the BGK equation with the collision invariants $\Phi(v) = (1, \xi, \frac{1}{2}\xi^2)^T$ and integrating over velocity space $d\xi$, one obtains the corrsponsing moments, eq. (8) to eq. (10).

$$\rho(x,t) = \int f d\xi \quad (8) \qquad \rho u(x,t) = \int \xi f d\xi \quad (9) \qquad E(x,t) = \int \frac{1}{2}\xi^2 f d\xi \quad (10)$$

2.1 Macroscopic features of Hydrodynamic and Rarefied Gas Flows in the SOD-Shock Tube



3 Deep learning

In this section deep learning with the focus on fully connected autoencoders and convolutional autoencoders will be introduced. Starting of with addressing important terminology whilst presenting the concept of autoencoders and continuing with the introduction of the fully connected and convolutional forwardpass , this section closes with ADAM [**?**]

an update to the backpropagation algorithm as well as important training methods.
The term deep learning situated around the much broader field of artificial intelligence stems from the use of deep feed forward networks also called multi layer perceptrons (MLP) or feed forward neural networks. Deep recurrent neural networks (RNN) are also used in this field but won't be covered in this thesis. In contrast to RNNs, information flows forward through these networks, which explains the name feed foward. Network refers to the typical composition of many different functions. In the following I will use the abbreviation MLP when talking about the aforementioned algorithm.
The task of any MLP is to approximate a function $f^\star(\mathbf{x};\mathbf{\Theta}) \approx f(\mathbf{x})$ through learning the values of the parameters $\mathbf{\Theta}$. As mentioned before $f^\star$ is a composition of functions eg. $f^3(f^2(f^1(\mathbf{x};\mathbf{\Theta}),\mathbf{\Theta}^2),\mathbf{\Theta}^3))$ in which f^1 denotes the input layer, f^3 the output layer and f^2 a hidden layer. The width of one hidden layer is referred to as the dimension of the vector valued input. The depth of the network describes the number of composed functions. In autoencoders the dimensions of input and output layer are identical.
Autoencoders are a special kind of neural network, that have a central hidden layer, that outputs a code c which should contain useful features of the input x while usually being of a lower dimension. Hereafter the code will be addressed as intrinsic variables highlighting the property of containing useful features of the input x. Autoencoders can be split in two parts, the encoder $h(x) = c$ which compresses the input and outputs the intrinsic variables c and the decoder $g(h) = \hat{x}$ which reconstructs the input from the intrinsic variables to output $\hat{y}$. The goal of autoencoders conflicts with the training objective. The former is to produce a code that describes the intrinsic features of the input, while the latter is to minimize the difference between x and $\hat{x}$. Therefore autoencoders need to be restrained from learning the identity function perfectly which in turn should drive the model to choose which instance to copy.
Obviously deep learning emphasizes the focus on the depth of a model. This is because linear models with just one layer can only approximate linear functions. Adding a non-linear activation function to the output of the proposed model wouldn't be sufficient in modelling any nonlinear behavior of the function. However the universal approximation theorem [**?**] states that MLPs with at least one hidden layer and any non-linear activation function can approximate any function given that enough hidden layers can be provided. In conclusion, MLPs are universal approximators.
Fully connected layers are called *Linear* in `PyTorch` because they compute a linear transformation of the input eq. (11), where x is the input vector, A is the weight matrix and b is a bias vector. The learnable parameters Θ are in this case the values in A and b. For a linear layer which takes a vector of size i as input and outputs a vector of size o there are $l = i \times o + o$ learnable parameters.

$$y = xA^T + b \tag{11}$$

4 Reduced Order Algorithms

4.1 Data Sampling

DIE ORIGINAL DTATENSTRUJTUR BESCHREIBEN For the autencoder using fully connected layers, the input vectors $y_o \in \mathbb{R}$ of size $n_{input} = n_\xi = 40$ are arranged in the

sampling matrix $S_{AE} \in \mathbb{R}^{5000x40}$ as seen in eq. (12) resulting in $n_S = 5000$ available samples. Note that the POD uses the same matrix transposed S_{AE}^T as input.

$$S_{AE} = \begin{bmatrix} f(\xi_1, t_1, x_1) & \cdots & f(\xi_n, t_1, x_1) \\ f(\xi_1, t_1, x_2) & \cdots & f(\xi_n, t_1, x_2) \\ \vdots & \vdots & \vdots \\ f(\xi_1, t_1, x_n) & \cdots & f(\xi_n, t_1, x_n) \\ f(\xi_1, t_2, x_1) & \cdots & f(\xi_n, t_2, x_1) \\ \vdots & \vdots & \vdots \\ f(\xi_1, t_n, x_n) & \cdots & f(\xi_n, t_n, x_n) \end{bmatrix} \tag{12}$$

$$S_{Conv} = \begin{bmatrix} n_{Filters} & f(\xi_1, \mathbf{t}, \mathbf{x}) \\ n_{Filters} & f(\xi_2, \mathbf{t}, \mathbf{x}) \\ \vdots & \\ n_{Filters} & f(\xi_n, \mathbf{t}, \mathbf{x}) \end{bmatrix} \tag{13}$$

Convolutional autoencoders use a different sampling matrix S_{Conv} due to their two dimensional capability resulting in $n_S = 40$ available samples eq. (13).$N_{Filters}$ varies over the succeeding layers, growing with the shrinkage of $(\mathbf{t}, \mathbf{x})$.

4.2 POD

The singular value decomposition of the input X [REF to Section 1] gives the optimal low-rank approximation $\tilde{X}$ of X eq. (14)[Eckard-Young].

$$\underset{\tilde{X}, s.t. rank(\tilde{X})=r}{\operatorname{argmin}} \quad ||X - \tilde{X}||_F = \tilde{U}\tilde{\Sigma}\tilde{V}^* \tag{14}$$

4.3 Autoencoders

Autoencoders have many hyperparameters determining their capability for compression and subsequent reconstruction. These parameters include : *number of layers*, *size of layers*, *activation functions*, *batch-size*, *learning rate*, *number of filters*, *stride width*, *kernel size*. Their finding is discussed in this section.
Originally the number of layers is determined, as they set the representational capacity of the model hence initiate over- and underfitting at an early stage. Figure 9 in section 5.4 shows the training and validation error calculated by:

$$Error = MSE/N_{Batches} \tag{15}$$

4.4 Architectures

4.4.1 Fully Connected

The autoencoder architecture 1.0 is a composition of five fully connected layers, eq. (16). The subscript F denotes fully connected layers opposed to the subscript C which denotes convolutional layers in section 4.4.2.

$$y_p = f_F^5(f_F^4(f_F^3(f_F^2(f_F^1(y_o))))) \tag{16}$$

Two input/output layers of size $n_{input} = 40$, two hidden layers of size $n_{hidden} = 20$ and one "bottleneck" layer of size $n_{code} = 3$. The trainingloss is the mean-squared error (MSE)

between the input and output of the autoencoder, where y_0 is the original inpput vector and y_p is the reconstructed output vector eq. (17).

$$MSE = \frac{(y_p - y_o)^2}{n_{batch}} \tag{17}$$

4.4.2 Convolutional

The convolutional autoencoder architecture 1.1 is a composition of six convolutional and three fully conected layers, eq. (18).

$$y_p = f_C^9(f_C^8(f_C^7(f_F^6(f_F^5(f_F^4(f_C^3(f_C^2(f_C^1(y_0)))))))))) \tag{18}$$

4.4.3 Training

During training every 1000 epochs a sample against its prediction was printed in order to link the value of the L1-Loss to a prediction. Using this method a first verification of the model was achieved. Continuing the search for any possible shortage of the models performance, that this method could not cover, eg. samples lying between every 1000 sample, that the model was not able to reconstruct correctly, a second verification process is conducted. Analysing the batch size for the architecture 1.0 the test errors in table 2 can be produced.

Kn	0.00001	Kn 0.01
Batch Size	L2-Error	
64	0.008	
32	0.0049	
16	0.0038	
8	0.0037	
4	0.0026	
2	0.0021	

Table 1: L2-Error over Batch-Size

4.5 Reduced Order Model

The compression of the input data y_0 yields a code $C \in \mathbb{R}^{ix5000}$, composed of the instrinsic variables c_i. The index i corresponds to the i-th intrinsic variable whereas their number is given by the input data. Each of them describes the transport of a discontinuity as seen in fig. 4. Hence the expolitability of the code in terms of constructing a ROM is not provided. On that account the method of characteristics [12] provides a means to bypass this shortage. It is necessary for $c_i(x,t)$ to satisfy the conservative condition eq. (19) and the transport equation eq. (20).

$$\frac{d}{dt}\int c_i \, dx = \frac{d}{dt} f_i = const. \tag{19}$$

$$\frac{\partial}{\partial t} c_i + \frac{\partial}{\partial x} f_i = 0 \tag{20}$$

The characteristics u_i describe the constant transport velocities for each variable c_i calculated using eq. (21). Subsequently enabling the usage of a simple plynomial interpoaltion

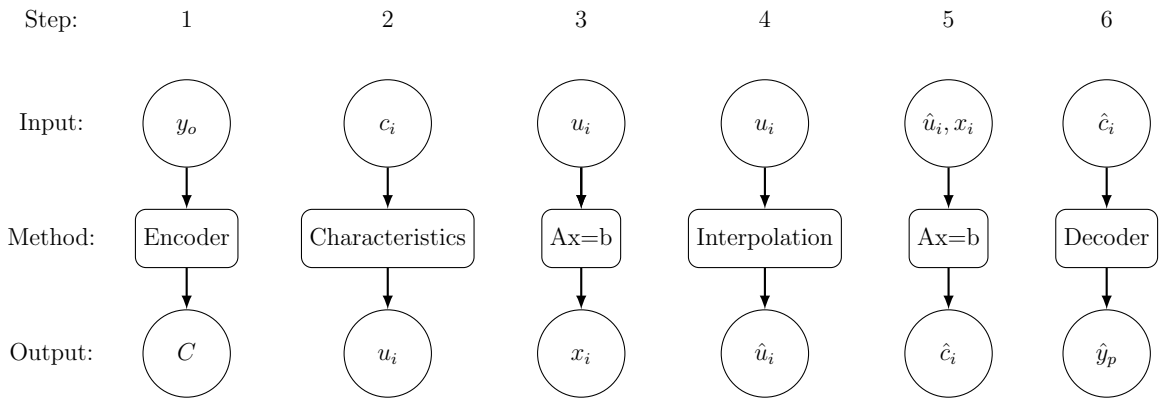


Figure 1: This figure shows the steps for obtaining a reduced oder model (ROM). Decoder and Encoder need to be used after training. In step one y_0 is the original input data, C is the Code. In step two c_i is the i-th intrinsic variable and u_i the correspnding characteristic. The eigenvalue problem in step 3 outputs x_i the eigenvector of A, a diagonal matrix composed of u_i and b is the corresponing i-th intrinsic variable c_i. In step 4 $\hat{u}_i$ is the interpolated vector to u_i. Step 5 solves the linear equation for the diagonal matrix A composed of $\hat{u}_i$ times the eigenvector x_i of the eigenvalueproblem in step 3. The output is $\hat{c}_i$ the i-th intrinsic variable corresponding to $\hat{u}_i$ the i-th interpolated characteristic.

of any degree. Furthermore a linear mapping $A_i x_i = c_i$ can be applied for the reconstruction of interpolated code variables $\hat{c}_i$. Figure 1 depicts this approach in detail.
Questions concering the capacity of this ROM, e.g. how many samples $\hat{n}_t$ are needed to reconstruct n_t timestamps, are analysed in section 5.

$$u_i = \frac{f_i(c_i^-) - f_i(c_i^+)}{c_i^- - c_i^+} \tag{21}$$

5 Results

5.1 Hydrodynamic Regime

In search for a reduced model of the BGK equation, a first reduction and analysis of the provided data in the hydrodynamic regime is conducted. The error over the L_2-Norm derived from eq. (24) assigns a value to each reduction algorithm enabling a evaluation. Furthermore the conservation quantities given in eq. (8) to eq. (10) of the prediction are

Algorithm	L_2
SVD	0.03
Fully Connected Autoencoder	0.002
Convolutional Autoencoder	0.02

Table 2: L2-Error over Batch-Size

analysed over the time average of each quantity. This normalization is given in eq. (23), where $\hat{\sigma}$ represents the given quantity. With more than 99% of the total cumultative energy S_N of the first five singular values calculated with eq. (22) the SVD provides an upper bound to the number of intrinsic features the autoencoder should extract. Figure 2 shows the singular values (left) and the cumulative energy (right).

$$S_N = \sum_{k=1}^{N} a_k \qquad \text{with a sequence} \qquad \{a_k\}_{k=1}^{n} \tag{22}$$

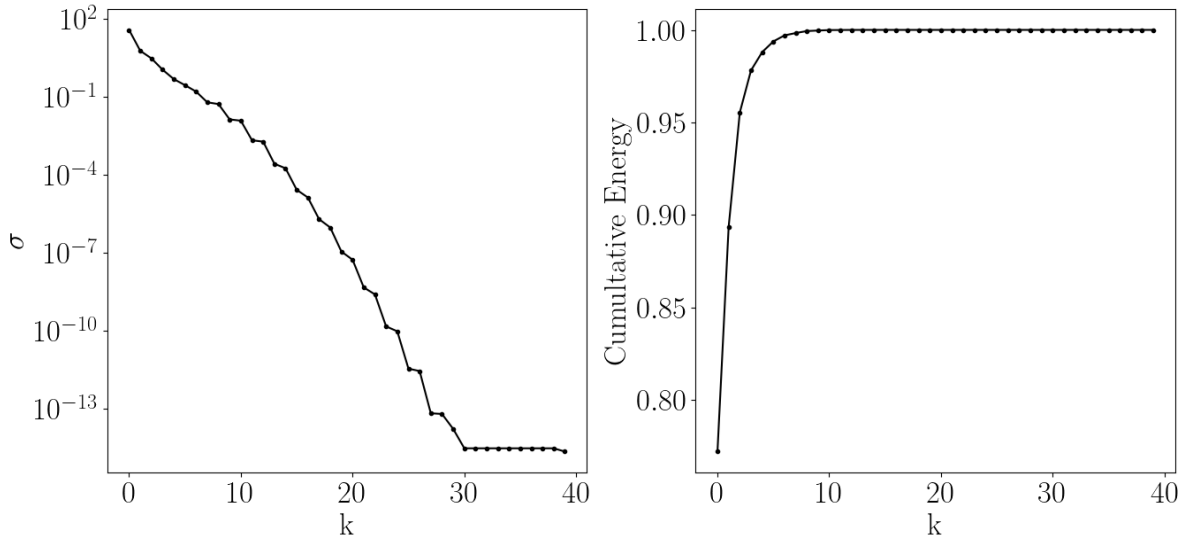


Figure 2: Singular Values (left) and cumultative enrgy (right) over the number of singular values

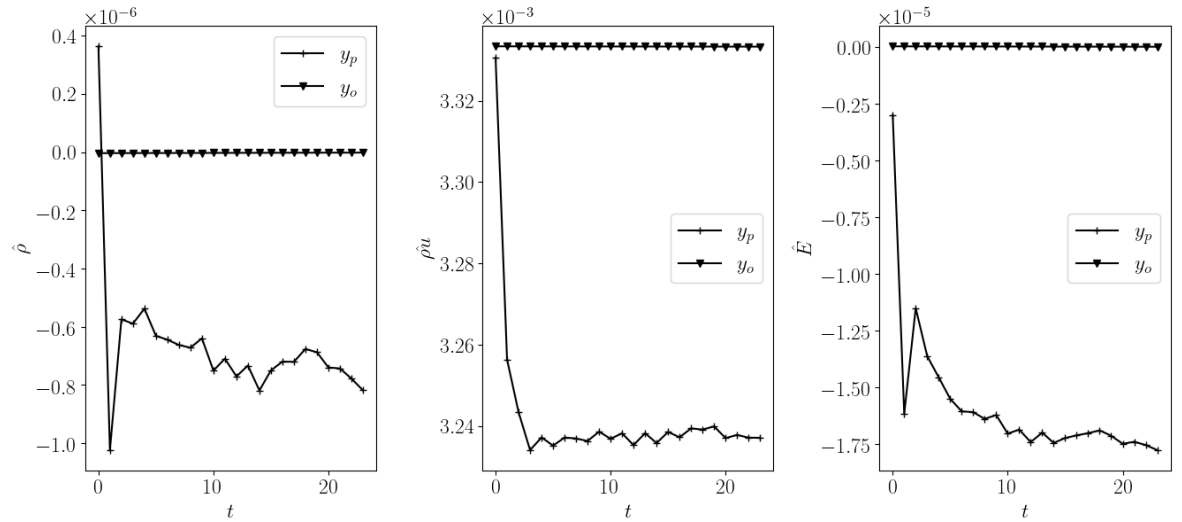
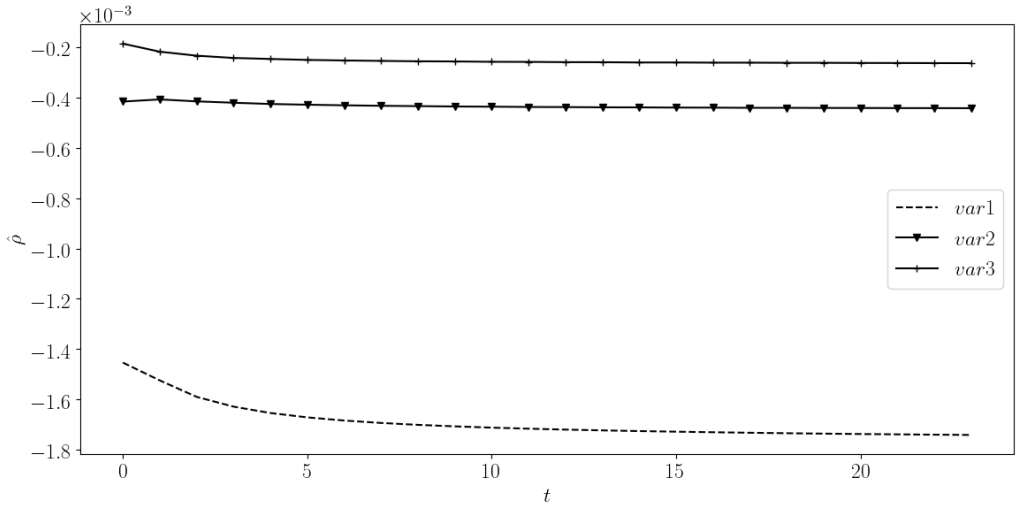


Figure 3: Normalized conservative quantities $\hat{\rho}$, $\hat{\rho u}$ and $\hat{E}$ as in eq. (23) for y_o and y_p.



$$\hat{\sigma} = \frac{\frac{d}{dt}\int \sigma \, dx}{\bar{\sigma}} = 0 \quad \text{and} \quad \bar{\sigma} = \frac{\iint \sigma \, dt dx}{\Delta t} \tag{23}$$

$$L_2 = \frac{||y_o - y_p||}{|||y_o|} \tag{24}$$

Given, that the autoencoder is able to achieve a reconstruction that is equal or below the threshold of the L_2-Norm form existing methods like SVD **??** and that conservation is preserved, a reduced order model can not be derived as described in **??**. Solely the the reconstuction can be validated. In addition the code needs to be a conservative system hence fulfilling eq. (8) to eq. (10).

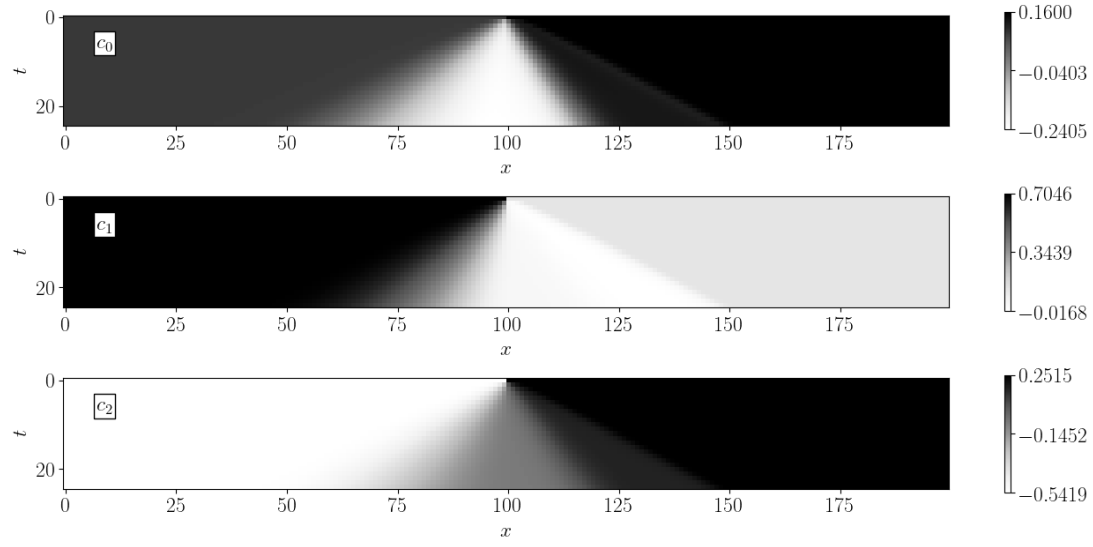


Figure 4: Code variable c_1,c_2 and c_3 over space x and time t of the fully connected autoencoder

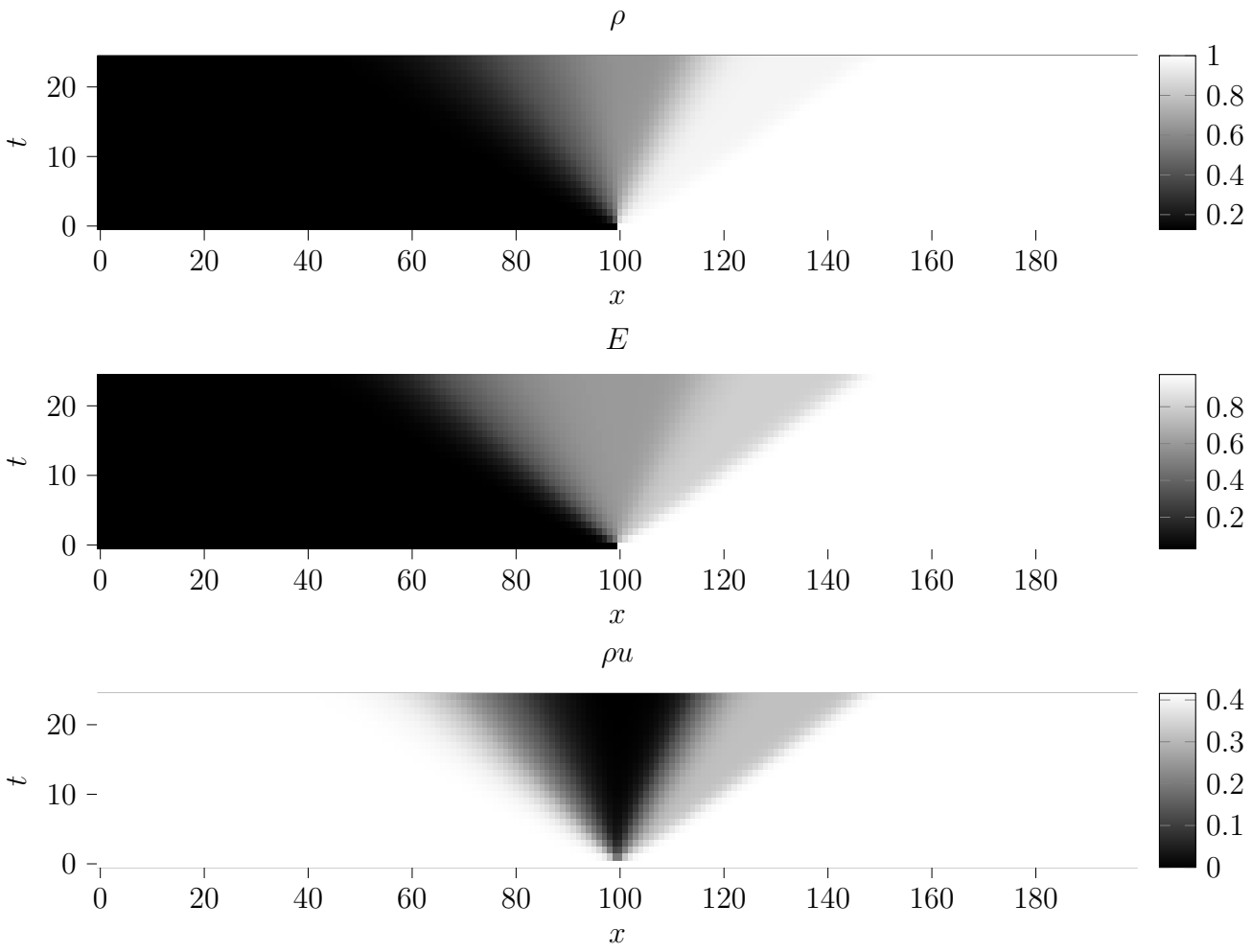


Figure 5: Macroscopic quantities ρ, E and ρ u of the original data.

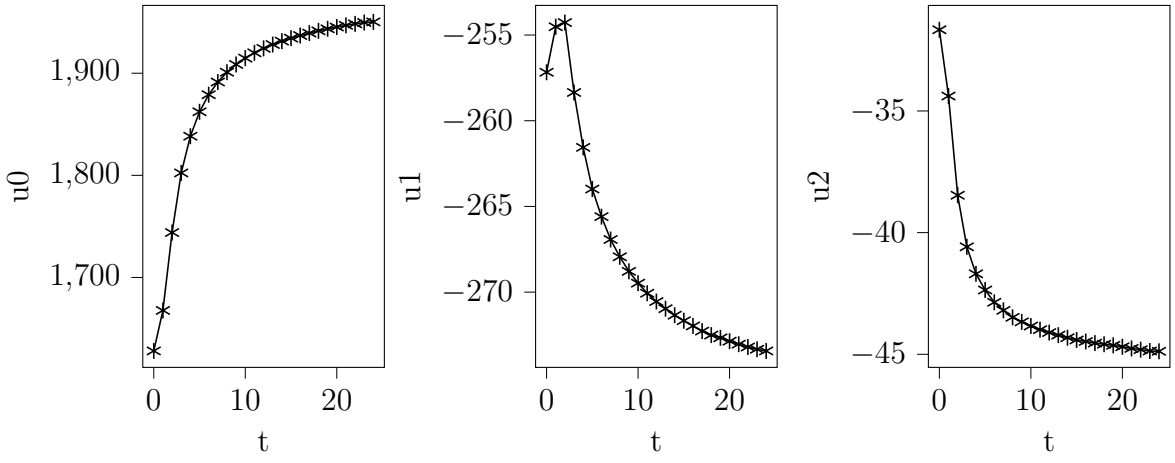


Figure 6: Characteristic velocities u0, u1, u2 of the code variables var0, var1, var2 respectively calculated as described in Section 4.5.

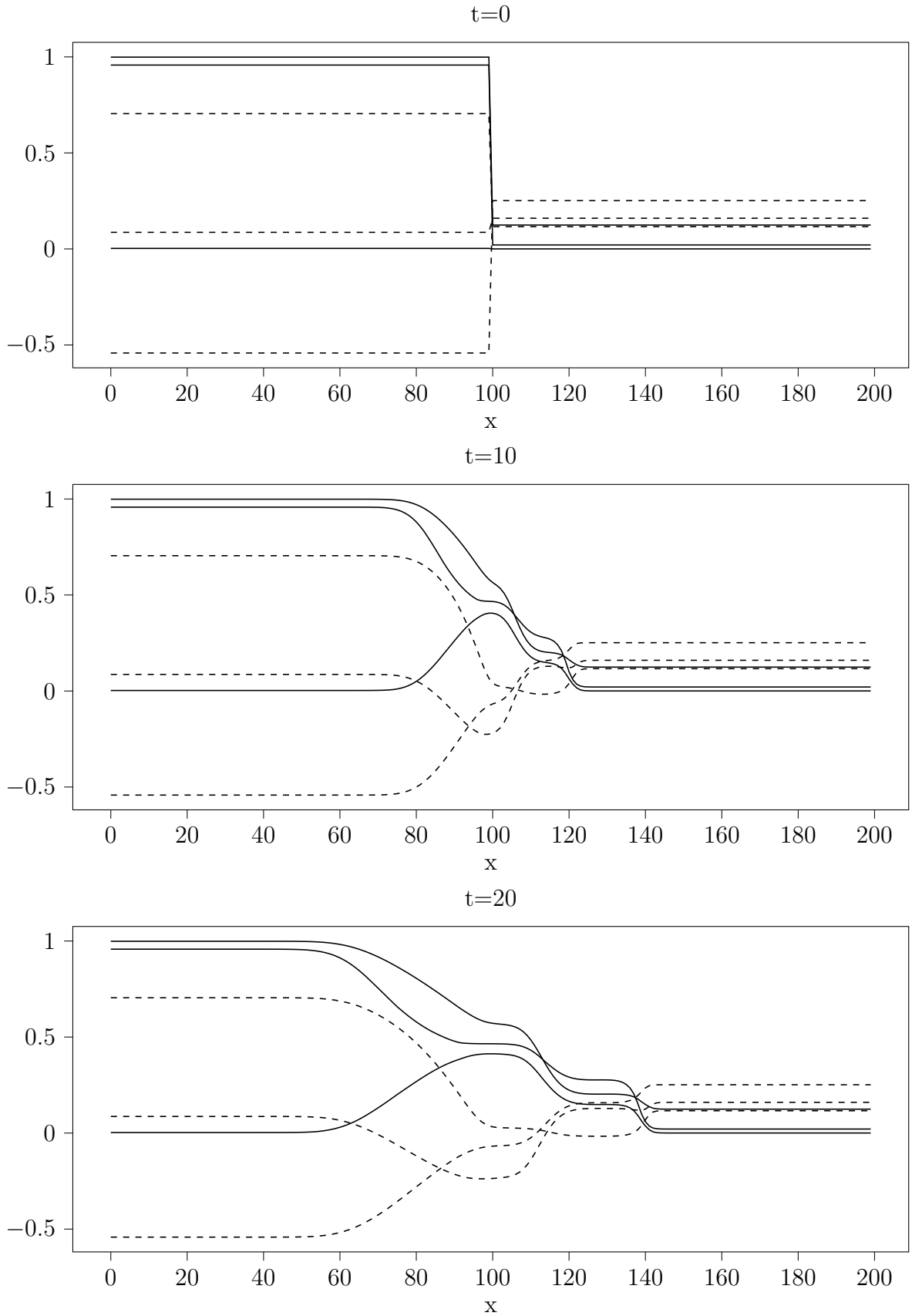


Figure 7: Code variables var0, var1, var3 (dashed lines - -) and macroscopic quantities ρ, E, ρu (full lines –) for three timestamps t_0, t_{10}, t_{20}.

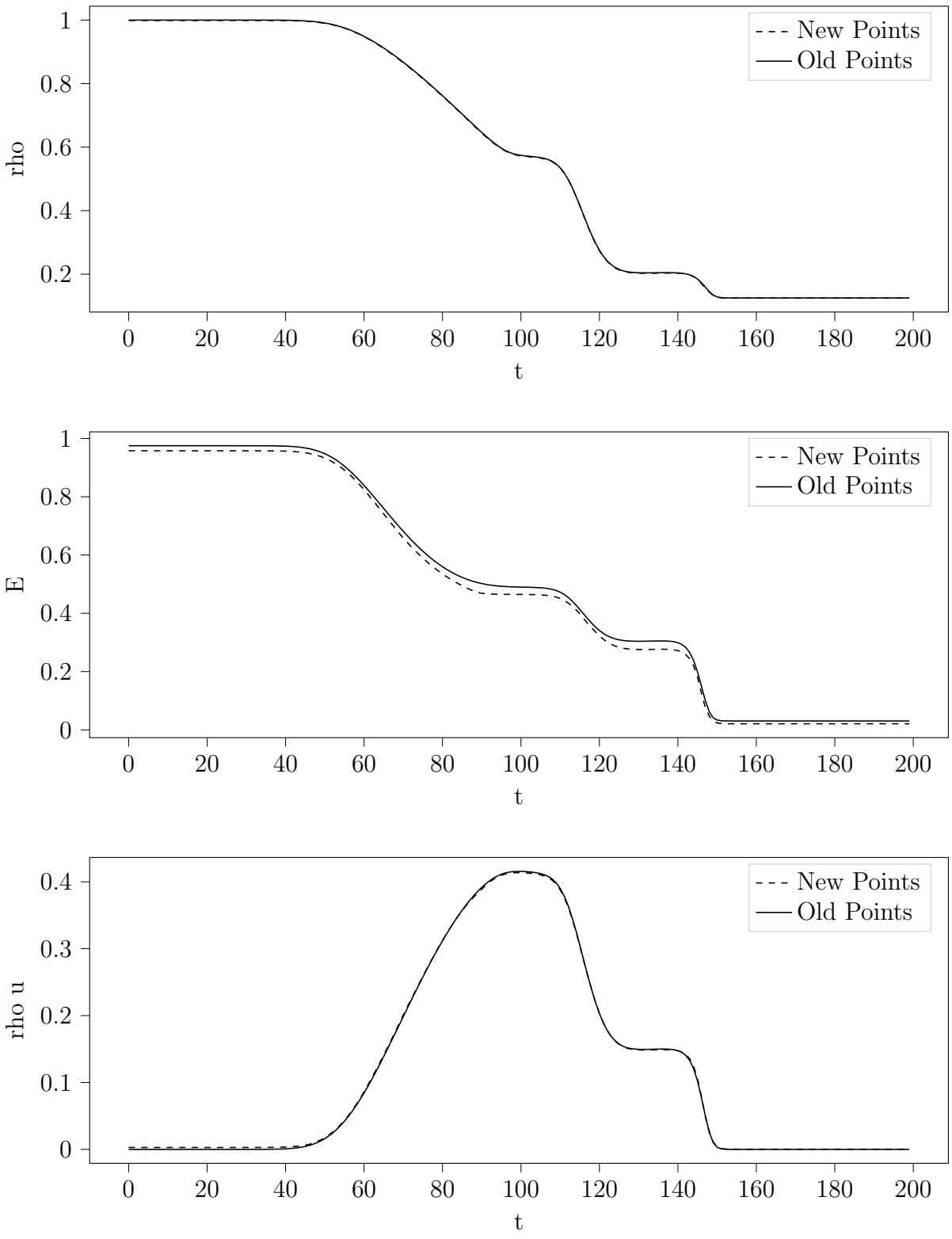


Figure 8: Resulting macroscopic quantities ρ, E, ρu after interpolating in time using the interpolation method described in 4.5. The interpolated quantity lies at timestamp t=24,5, while the original quantity lies at timestamp t=25.

5.2 Rarefied Regime

5.3 Discussion and Outlook

References

[1] Bhatnagar, Gross, and Krook. A model for collision processes in gases. 1954.

[2] Thomas Franz. *Reduced-order modeling of steady transonic flows via manifold learning.* 2016.

[3] Steve L. Brunton and J. Nathan Kutz. *Data driven science and engineering.* 2019.

[4] Florian Bernard, Angelo Iollo, and Sebatian Riffaud. Reduced-order model for the bgk equation based on pod and optimal transport. 2018.

[5] Kookjin Lee and Kevin T. Carlberg. Model reduction of dynamical systems on nonlinear manifolds using deep convolutional autoencoders. 2019.

[6] Ian J. Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning.* MIT Press, Cambridge, MA, USA, 2016. `http://www.deeplearningbook.org`.

[7] D.E. Rumelhart, G.E. Hinton, and R.J. Williams. Learning internal representations by error propagation. 1986.

[8] Dana H. Ballard. Modular learning in neural networks. 1987.

[9] Salah Rifai, Pascal Vincent, Xavier Muller, Xavier Glorot, and Yoshua Bengio. Contractive auto-encoders: Explicit invariance during feature extraction. 2011.

[10] Salah Rifai, Yann N Dauphin, Pascal Vincent, Yoshua Bengio, and Xavier Muller. *The Manifold Tangent Classifier.* Curran Associates, Inc., 2011.

[11] Salah Rifai, Yoshua Bengio, Yann Dauphin, and Pascal Vincent. A generative process for sampling contractive auto-encoders, 2012.

[12] Hervé Le Dret and Brigitte Lucquin. *Partial Differential Equations: Modeling, Analysis and Numerical Approximation* -. Birkhäuser, Basel, 2016.

5.4 Appendix A

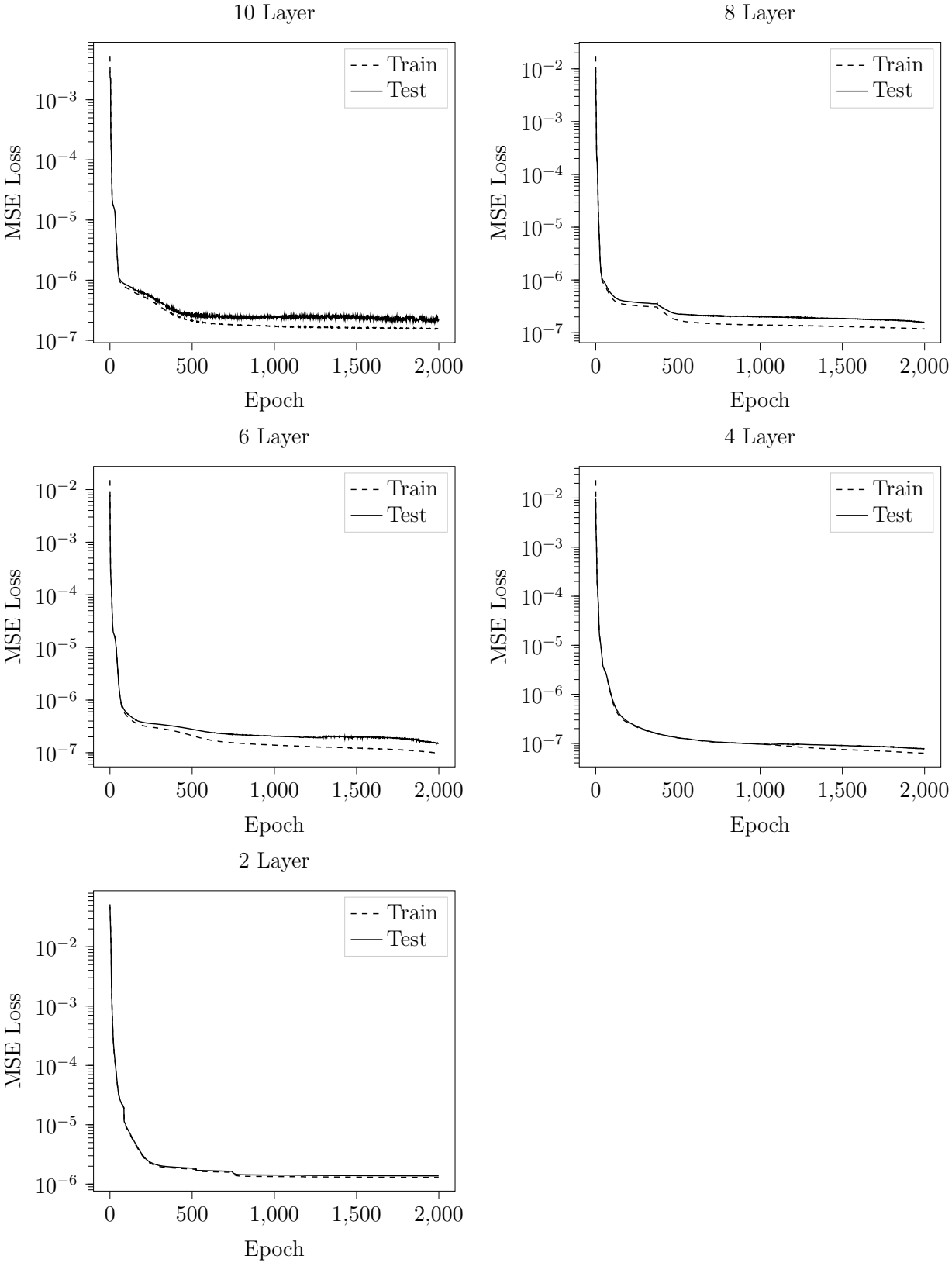


Figure 9: Analysis of layer size for five different architectures, showing the error for training and testing.