



# Application of Koopman operator for model-based control of fracture propagation and proppant transport in hydraulic fracturing operation



Abhinav Narasingam, Joseph Sang-Il Kwon \*

Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, TX 77845, USA  
Texas A&M Energy Institute, Texas A&M University, College Station, TX 77845, USA

## ARTICLE INFO

### Article history:

Received 28 October 2019  
Received in revised form 19 February 2020  
Accepted 10 May 2020  
Available online 22 May 2020

### Keywords:

Koopman operator  
Model predictive control  
Distributed parameter systems  
Moving boundary problem  
Hydraulic fracturing  
Extended dynamic mode decomposition

## ABSTRACT

This work explores the application of the recently developed Koopman operator approach for model identification and feedback control of a hydraulic fracturing process. Controlling fracture propagation and proppant transport with precision is a challenge due in large part to the difficulty of constructing approximate models that accurately capture the characteristic moving boundary and highly-coupled dynamics exhibited by the process. Koopman operator theory is particularly attractive here as it offers a way to explicitly construct linear representations for even highly nonlinear dynamics. The method is data-driven and relies on lifting the states to an infinite-dimensional space of functions called observables where the dynamics are governed by a linear Koopman operator. This work considers two problems: (a) fracture geometry control, and (b) proppant concentration control. In both cases, an approximate linear model of the corresponding dynamics is constructed and used to design a model predictive controller (MPC). The manuscript shows that in the case of highly nonlinear dynamics, as observed in the proppant concentration, use of canonical functions in the observable basis fails. In such cases, a priori system knowledge can be leveraged to choose the required basis. The numerical experiments demonstrate that the Koopman linear model shows excellent agreement with the real system and successfully achieves the desired target values maximizing the oil and gas productivity. Additionally, due to its linear structure, the Koopman models allow convex MPC formulations that avoid any issues associated with nonlinear optimization.

Published by Elsevier Ltd.

## 1. Introduction

Over the past two decades, the natural gas from shale plays has been one of the fastest growing total primary energy contributors in the United States; it has risen from less than 1% of domestic gas production in 2000 to over 20% by 2010 [1]. Although shale gas has been produced for years from shales with natural fractures, the recent shale boom has been primarily due to the use of two technologies: horizontal drilling and hydraulic fracturing. Typically, shale formations are characterized by low matrix permeabilities ( $\leq 0.01 - 0.0001$  mD) and hence preclude the use of conventional drilling techniques. Therefore, the recovery of shale oil and gas in commercial quantities is economically viable by creating extensive artificial fractures around wellbores [2]. Specifically, in the fracturing process, a mixture of water, sand and chemicals is injected into the horizontal borehole of the well at very high pressure to fracture the shale rocks and release the gas. The created fractures facilitate the extraction of

oil and gas by providing high conductivity pathways from the formation to the wellbore while the proppant (sand whose permeability is very high compared to the surrounding rock) trapped inside the fracture walls at the end of the process increases the formation permeability within the stimulated reservoir volume. Thus, the final fracture conductivity is paramount in maximizing the productivity of the stimulated well, which is the ultimate goal of a fracturing process in practice.

The final fracture conductivity chiefly depends on two factors: propped fracture geometry and proppant distribution throughout the fracture at the end of pumping. In order to achieve a high fracture conductivity, it is essential to create fractures with the desired geometry as well as achieve uniform proppant concentration throughout the fracture at the end of pumping operation. Traditionally, during the planning stage, an optimization-based methodology called Unified Fracture Design (UFD) is used for the design of optimal hydraulic fracturing treatments [3,4]. Specifically, based on the total quantity and properties of the proppant to be injected, reservoir properties and drainage area, UFD determines the optimal fracture dimensions such as the fracture half-length and fracture width that maximizes the well productivity index. Once, a desired geometry is decided, during the

\* Corresponding author at: Texas A&M Energy Institute, Texas A&M University, College Station, TX 77845, USA.  
E-mail address: [kwonx075@tamu.edu](mailto:kwonx075@tamu.edu) (J.S. Kwon).

implementation stage, the fracture geometry is adjusted by varying the flow rate of the injected proppant. Simultaneously, the spatial distribution of the proppant concentration across the fracture is regulated by varying the inlet proppant concentration (i.e., pumping schedule). However, the problem with controlling these petroleum systems is that the mathematical models used to describe them are very complex due to highly-coupled nonlinear system dynamics and parametric uncertainty. Solving these forward problems at very fine scales is computationally very expensive and becomes prohibitive in the case of (optimal) controller design where such forward problems need to be solved iteratively.

Owing to this, initial efforts viewed the fracturing process as an open-loop problem where the pumping schedule followed a pre-determined power-law which assumed a constant leak-off and does not take into account practical constraints [5]. To deal with its limitations, researchers also leveraged exhaustive forward simulations which were used to adjust the pumping schedule iteratively until the desired fracture geometry and fracture conductivity were achieved. This approach outperforms Nolte's pumping schedule at the expense of computational resources [6, 7]. In recent years, several works successfully considered the closed-loop operation of the hydraulic fracturing process by using model-based predictive controllers that take advantage of feedback measurements to compensate for minor model uncertainty and disturbances [8–10]. The performance of these real-time control schemes highly depends on the availability of computationally efficient approximate models capable of capturing the nonlinear process behavior. However, a nonlinear trajectory can explore many regions of the state space, and therefore, different portions of such a trajectory can be associated with distinct regimes. Moreover, the hydraulic fracturing process is characterized by a moving boundary; the fracture propagates in the horizontal direction with time. In such systems, it has been previously shown that traditional model approximation methods may fail to accurately capture the dynamics [11–13]. This was attributed to the fact that the spatial domain of interest as well as the dominant spatial pattern of the system changes with time. To deal with this, the research works mentioned above attempted to tailor the models to capture various portions of the system's trajectory using multiple temporally local models. However, obtaining optimal clustering of temporal domain requires solving a large scale mixed integer nonlinear programming problem which is computationally expensive, although it is solved offline. Additional efforts have been made to enlarge their range of applicability [14]. Nevertheless, these temporally local models suffer from generalizability on the entire state space; in other words, they predict poorly when used in the operating regimes that are significantly different from the training ones.

In this work, we investigate an alternative approach where we employ Koopman operator theory to generate data-driven models that approximate the hydraulic fracturing dynamics. It is particularly attractive because of its ability to provide (nearly) globally valid linear models. To be more specific, the Koopman linear models are often valid in a larger domain unlike local models whose validity is, in general, limited by the training data [15]. It was first introduced by B. O. Koopman in 1931 [16], where he showed that any finite-dimensional nonlinear system has an equivalent infinite-dimensional linear representation in the space of (all possible) functions of the system states. In this function space, the evolution of the system is governed by an infinite-dimensional linear operator. This operator theoretic approach is fundamentally different from classical approaches in the sense that instead of using the state vectors in a state space, it lifts the system (nonlinear transformation) to the space of observables (measures) using a basis of nonlinear functions and is analyzed

there. In addition to the system evolution being linear in the lifted space, the operator theoretic approach facilitates the data-driven analysis of dynamical systems. In particular, recent advances in numerical techniques such as generalized Fourier and Laplace analysis [17], Dynamic Mode Decomposition (DMD) [18,19] and Extended Dynamic Mode Decomposition (EDMD) [20] have been successful in efficiently computing finite-dimensional approximations to the Koopman operator from time-series data, obtained from simulations or experiments. Since its revival by Mezić [21], there has been an increased interest in the Koopman operator approach, and it has seen practical implementation in a broad range of applications [22–25]. Apart from system identification, the Koopman operator has been successfully applied in the context of nonlinear stability analysis [26,27], state estimation [28–30] and control of nonlinear dynamical systems [31–40].

Our contribution in this work is demonstrating, on a non-trivial engineering application of hydraulic fracturing, the data-driven Koopman method for constructing linear models that can be readily used within a predictive control framework to maximize the productivity of a fractured well. To do so, we first apply the Koopman-based system identification method to create a dynamic model of the fracturing process and verify that it captures the system's true dynamic behavior. Then, we design a model predictive controller (MPC) using the model obtained in the previous step. Since the developed models are linear in the space of observable functions, the predictive controller when formulated in this space will lead to a standard convex optimization problem which allows for extremely fast solutions compared to the original nonlinear problem. We consider two case studies of varying complexities: first, we apply the approach to regulate fracture geometry by considering only the fracture propagation dynamics governed by a partial differential equation (PDE). As a more sophisticated case study, we consider the simultaneous fracture propagation and proppant transport to determine the optimal pumping schedule that achieves uniform proppant concentration throughout the fracture at the end of operation. The overall dynamic model describing rock deformation, fluid flow, and proppant transport is a set of highly-coupled nonlinear equations defined over a time-dependent spatial domain making it considerably more complex than fracture propagation alone. Therefore, this calls for careful curating of the Koopman basis and we show how a priori system knowledge can be incorporated during system identification in choosing appropriate observable functions.

The rest of this paper is organized as follows: In Section 2, we present some preliminaries on the Koopman operator and describe its utilization for model identification and MPC design. In Section 3, we briefly describe the dynamic model representing the simultaneous fracture propagation and proppant transport during a hydraulic fracturing process. Section 4 discusses the application of Koopman operator theory to the hydraulic fracturing process. Specifically, two different control problems with different complexities are presented. The results of applying the Koopman system identification technique to the simulated data are summarized and the performance of the model predictive controllers is studied. In Section 5, concluding remarks and perspectives are provided.

## 2. Preliminaries

In this section, we present some preliminaries on the Koopman operator theory and its application for model identification and controller synthesis.

## 2.1. Koopman operator

Consider a discrete controlled nonlinear dynamical system given below:

$$x_{k+1} = F(x_k, u_k) \quad (1)$$

where  $x_k \in \mathcal{X} \subseteq \mathbb{R}^n$  is the state of the system,  $u_k \in \mathcal{U} \subseteq \mathbb{R}^m$  is the input,  $F$  is the flow map that evolves the system states forward in time, and  $k \in \mathbb{Z}$  represents the time step such that  $t_k = k\Delta$  where  $\Delta$  is the sampling time.

Let us define a set of scalar-valued *observables* that are functions of the states and the inputs,  $g : \mathcal{X} \times \mathcal{U} \rightarrow \mathbb{R}$ . Each observable is an element of an infinite-dimensional function space  $\mathcal{G}$  which, for example, can be defined by the Lebesgue square-integrable functions,  $\mathcal{G} = \ell^2(\mathcal{X} \times \mathcal{U}, \mathbb{R})$ , or other appropriate spaces [41]. In this infinite-dimensional function space, the flow of the system is governed by the Koopman operator,  $\mathcal{K} : \mathcal{G} \rightarrow \mathcal{G}$  that defines the dynamics of observables  $g \in \mathcal{G}$  along the trajectories of the system as:

$$\mathcal{K}g \triangleq g \circ F \quad (2)$$

By definition, the Koopman operator is linear even though the underlying dynamical system is nonlinear. For all  $g_1, g_2 \in \mathcal{G}$  and all  $\alpha, \beta \in \mathbb{R}$ , it satisfies

$$\begin{aligned} \mathcal{K}(\alpha g_1 + \beta g_2) &= (\alpha g_1 + \beta g_2) \circ F \\ &= (\alpha g_1 \circ F) + (\beta g_2 \circ F) \\ &= \alpha \mathcal{K}g_1 + \beta \mathcal{K}g_2 \end{aligned} \quad (3)$$

The system identification method described below exploits the fact that any finite-dimensional nonlinear system can be equivalently represented using an infinite-dimensional linear system by transforming the traditional state space to the space of functions (observables) of the system's states and inputs. Please note that the Koopman operator has originally been proposed for autonomous dynamical systems, and it can be adopted to controlled systems by considering joint observations of state and input, i.e.,  $\mathcal{K}$  for the above system Eq. (1) can be considered as the classical Koopman operator for the augmented system,  $x^+ = F(x, u)$  and  $u^+ = \mathcal{S}(u)$  where  $\mathcal{S}(u)$  indicates a forward shift operator of a known input signal. This is just one way to adopt the Koopman operator to controlled systems, and for more generalizations the readers can refer to [31,42,43].

## 2.2. System identification using EDMD

The Koopman operator theory is conceptually developed on the infinite-dimensional function space  $\mathcal{G}$ . However, it is not practically feasible unless we can determine finite-dimensional approximations to the Koopman operator without a great loss in accuracy. To do this, consider a finite-dimensional subspace  $\bar{\mathcal{G}} \subset \mathcal{G}$  spanned by a set of basis functions  $\phi(x, u) = [\phi_1(x, u), \phi_2(x, u), \dots, \phi_{N_\phi}(x, u)]^T$ . Now, any observable function  $\psi \in \bar{\mathcal{G}}$  can be represented as a linear combination of these basis functions as follows:

$$\psi = c_1\phi_1 + c_2\phi_2 + \dots + c_{N_\phi}\phi_{N_\phi} = c^T\phi \quad (4)$$

For these functions in  $\bar{\mathcal{G}}$ , we seek to generate the finite-dimensional approximation of the Koopman operator, denoted as  $K \in \mathbb{R}^{N_\phi \times N_\phi}$ . Because, typically  $\bar{\mathcal{G}}$  is not invariant with respect to  $\mathcal{K}$ , there is a residual term which is minimized in the  $L_2$ -norm sense via linear regression [44]. Intuitively, if these observables  $\psi$  represent the system's outputs (or equivalently, sensor measurements), a cost function to be minimized in an optimal control problem, or nonlinear constraints, then the evolution of these "nonlinear" observables can now be analyzed using a linear

system via  $K$ . This is done by carrying a nonlinear transformation of the system states (outputs) to the so-called "lifted" space using the set of  $N_\phi$  basis functions.

In its most general form, the Koopman linear system is presented as follows:

$$\phi(x_{k+1}, u_{k+1}) = K(x_k, u_k)\phi(x_k, u_k) \quad (5)$$

where  $\phi \in \mathbb{R}^{N_\phi}$  and usually  $N_\phi \gg n$ . For the augmented state  $(x, u)$ , the above formulation will admit any nonlinear dynamical system. However, as the dimension of the original system increases, the number of basis functions to be considered can grow to infeasible levels. So, to make the computations tractable and amenable in the case of predictive controller design (please see Section 2.3 for more details), the simplifications described below are introduced in the structure of the nonlinear transformation [34].

Specifically, the objective is to identify a linear dynamical system of the following form (similar to a linear time invariant state space model) in the space of observables using time-series data generated by the control system Eq. (1).

$$\phi(x_{k+1}) = A\phi(x_k) + Bu_k \quad (6)$$

For this purpose, we use the recently developed EDMD algorithm. To construct a finite dimensional approximation to the Koopman operator for the controlled system in Eq. (1), the EDMD algorithm requires

1. a time-series data set of  $N_t$  snapshot pairs satisfying the dynamical system in Eq. (1) which can be organized in the following matrices.

$$\begin{aligned} X &= [x_1, x_2, \dots, x_{N_t}], \quad Y = [y_1, y_2, \dots, y_{N_t}], \\ U &= [u_1, u_2, \dots, u_{N_t}] \end{aligned} \quad (7)$$

Note that we use  $y_k$  instead of  $x_{k+1}$  here because the data above need not be temporally ordered as long as it satisfies  $y_k = F(x_k, u_k)$ .

2. a library of nonlinear basis functions  $\{\phi_1, \phi_2, \dots, \phi_{N_\phi}\}$  whose span is  $\bar{\mathcal{G}} \subset \mathcal{G}$ .

The EDMD algorithm then seeks to solve a least-squares problem to obtain  $K$  which is the transpose of the finite-dimensional approximation to the Koopman operator,  $\mathcal{K}$ :

$$\min_K \sum_{k=1}^{N_t} \|\phi(y_k, u_k) - K\phi(x_k, u_k)\|_2^2 \quad (8)$$

In order to obtain the form described in Eq. (6), the following simplification can be introduced into the formulation. Since we are not interested in predicting the future values of inputs, without loss of generality we can assume that

$$\phi(x, u) = \begin{bmatrix} \phi(x) \\ u \end{bmatrix}, \quad \phi(y, u) = \begin{bmatrix} \phi(y) \\ u \end{bmatrix} \quad (9)$$

which is to say that the nonlinear observable functions are applied to the system states alone and not the inputs. Suppose that  $N$  nonlinear functions are used to lift the states to the observable space and the  $m$  system inputs are added separately to the basis as shown in Eq. (9), we have  $N_\phi = N + m$ . To obtain the evolution of lifted states, we can then disregard the last  $m$  components of each of the terms in Eq. (8), decompose the Koopman matrix as  $K = [A, B]$  and use only the first  $N$  rows in  $K$ , which leads to the following minimization problem

$$\min_{A, B} \sum_{k=1}^{N_t} \|\phi(y_k) - A\phi(x_k) - Bu_k\|_2^2 \quad (10)$$

The practical solution to the above equation is obtained by using regularization via a truncated singular value decomposition and the value of  $K$  that minimizes Eq. (10) is given by

$$K_N := [A, B] = \phi_{XY} \phi_{XX}^\dagger \quad (11)$$

where  $^\dagger$  denotes the pseudoinverse, and the matrices are computed as

$$\phi_{XX} = \begin{bmatrix} \phi_X \\ U \end{bmatrix} \begin{bmatrix} \phi_X \\ U \end{bmatrix}^T, \quad \phi_{XY} = \phi_Y \begin{bmatrix} \phi_X \\ U \end{bmatrix}^T \quad (12)$$

$$\phi_X = [\phi(x_1), \dots, \phi(x_{N_t})], \quad \phi_Y = [\phi(y_1), \dots, \phi(y_{N_t})]$$

with

$$\phi(x_k) = \begin{bmatrix} \phi_1(x_k) \\ \vdots \\ \phi_N(x_k) \end{bmatrix} \quad (13)$$

Please note that any solution to Eq. (11) is a solution to Eq. (10), and the formulation of Eq. (11) has an advantage of being independent of the number of data samples  $N_t$ . Therefore, the Koopman linear system obtained using the above algorithm will be in the form of the controlled linear dynamical system given by Eq. (6), where  $\phi \in \mathbb{R}^N$  is the lifted state in the observable space, and  $A \in \mathbb{R}^{N \times N}$  and  $B \in \mathbb{R}^{N \times m}$  are the matrices that describe the system dynamics in the lifted space.

The Koopman linear system Eq. (6), derived using the above algorithm, governs the evolution of the basis  $\phi$  in the lifted space. With the use of these basis functions, all observables of interest  $\psi \in \mathbb{R}^{n_\psi}$  can be determined (refer to Eq. (4)) by simply using a matrix of coefficients as  $\psi(x_k) = C\phi(x_k)$ , where  $C \in \mathbb{R}^{n_\psi \times N}$ . In summary, the basis functions are first evolved linearly using the Koopman operator, and the value of (required) observable is then computed by linearly combining these basis functions. In reality, not all observables can be contained in the span of the chosen subspace (recall, we are only using a finite-dimensional truncated function space). Thus, the matrix  $C$  must be chosen such that the projection of  $\psi$  onto the span $\{\phi_1, \dots, \phi_N\}$  is minimized in the  $L_2$ -norm sense as below:

$$\min_C \sum_{k=1}^{N_t} \|\psi(x_k) - C\phi(x_k)\|_2^2 \quad (14)$$

When dealing with dynamical systems, since the goal is stability analysis or controller synthesis, it is advantageous to reproduce the state dynamics. Therefore, the observable of interest here is the state itself,  $\psi = x$ . In such cases, we can assume that the basis also contains the state observable, i.e.,  $[\phi_1, \dots, \phi_n]^T = x$ . In particular, such a set of basis functions is said to be *state-inclusive*, and typically the solution to matrix  $C$  in Eq. (14) is trivial and can be obtained by  $C = [I_n, 0_{N-n}]$ .

### 2.3. Koopman model predictive control

In this section, we briefly describe how the Koopman linear model obtained in Eq. (6) is used within an MPC scheme. In an MPC controller, an optimization problem is solved repeatedly over finite prediction horizons with respect to control inputs and predicted outputs of the system and a feedback behavior is achieved by taking process measurements as the initial condition. Typically, for nonlinear systems this is a nonconvex optimization problem due to the nonlinear dynamics. For the system in Eq. (1),

the MPC problem at sampling time step  $k$  is given by

$$\begin{aligned} & \text{Minimize}_{u_k, \dots, u_{k+N_p-1} \in \mathcal{U}} \sum_{i=1}^{N_p-1} x_{k+i}^T Q x_{k+i} + u_{k+i}^T R u_{k+i} + x_{k+N_p}^T Q_{N_p} x_{k+N_p} \\ & \text{s.t.} \quad x_{k+i} = F(x_{k+i-1}, u_{k+i-1}), \quad i = 1, \dots, N_p \\ & \quad x_k = x_k^{\text{measured}} \\ & \quad c(x_{k+i}) \leq b, \quad i = 1, \dots, N_p \\ & \quad \underline{u} \leq u \leq \bar{u} \end{aligned} \quad (15)$$

where the cost matrices  $Q, Q_{N_p} \in \mathbb{R}^{n \times n}$  and  $R \in \mathbb{R}^{m \times m}$  are positive semi-definite matrices,  $N_p$  is the prediction horizon,  $x_k^{\text{measured}}$  is the state measured at time  $t_k = k\Delta$  where  $\Delta$  is the sampling time during the closed-loop operation,  $c: \mathbb{R}^n \rightarrow \mathbb{R}^{n_c}$  are nonlinear functions used to define the state constraints,  $b \in \mathbb{R}^{n_c}$  is a constant vector and  $\underline{u}$  and  $\bar{u}$  denote the lower and upper bounds, respectively, on the manipulated inputs.

In contrast to the above nonlinear optimization problem, in the Koopman MPC framework a convex quadratic optimization problem (QP) is solved instead via lifting to the observable functions space. At time step  $k$ , the predictions of the system trajectory are initialized from the lifted state  $\phi_k = \phi(x_k)$ . Similarly, the objective function and the state constraints are all transformed to the lifted space. In this lifted space, the original nonlinear equation is replaced by the linear dynamics obtained in Eq. (6). Additionally, any nonlinear constraints associated with the system can be absorbed in the basis and therefore predicted in a linear manner through Eq. (6). For simplicity, if we consider the following output mapping  $\phi$

$$\phi = \begin{bmatrix} \tilde{\phi}(x) \\ c(x) \end{bmatrix} \quad (16)$$

where  $\tilde{\phi} \in \mathbb{R}^{N-n_c}$  are some nonlinear functions in the lifted space, the nonlinear state constraints in  $x$  then translate to linear constraints in  $\phi$ , i.e.,  $c(x) = E\phi(x)$  where  $E = [0_n \quad I_{n_c}]^T$ . Based on the above transformations, the Koopman MPC, in discrete formulation, solves the following QP problem at time step  $k$  of the closed-loop operation.

$$\begin{aligned} & \text{Minimize}_{\phi_k, \dots, \phi_{k+N_p-1} \in \mathcal{U}} \sum_{i=1}^{N_p-1} \phi_{k+i}^T Q \phi_{k+i} + u_{k+i}^T R u_{k+i} + \phi_{k+N_p}^T Q_{N_p} \phi_{k+N_p} \\ & \text{s.t.} \quad \phi_{k+i} = A\phi_{k+i-1} + Bu_{k+i-1}, \quad i = 1, \dots, N_p \\ & \quad \phi_k = \phi(x_k^{\text{measured}}) \\ & \quad E\phi_{k+i} \leq b, \quad i = 1, \dots, N_p \\ & \quad \underline{u} \leq u \leq \bar{u} \end{aligned} \quad (17)$$

This feature of incorporating all nonlinearities within the observable basis is one of the main attractions of the Koopman operator approach. For example, in our previous work, the same idea was used to propose a stabilizing Lyapunov-based MPC formulation for nonlinear systems [40]. It can be further extended to translate any nonlinear objective functions such as those observed in Economic MPC formulations [45] to convex quadratic functions.

**Remark 1.** It is emphasized that the convex MPC presented in Eq. (17) is only an approximation of the original nonlinear MPC problem in Eq. (15). But, nonlinear MPC problems are typically  $np$ -hard whereas the presented convex QP is polynomial time solvable and as long as the predictions are accurate, we expect the solution of the linear MPC problem to be close to the optimal solution of the nonlinear MPC problem. Please note that the rigorous quantification of approximation errors in the Koopman

operator is still an actively researched area with several results available for the convergence of the EDMD algorithm under some assumptions [44,46,47].

**Remark 2.** Regarding the stability of the closed-loop system, the Koopman methodology presents an exciting avenue to develop stabilizing feedback controllers. Specifically, the Lyapunov theory can be naturally extended to analyze the stability of the resulting closed-loop system. Several successful results exploring this can be found in the following works [26,39,40]. One idea is to design a stabilizing control framework by explicitly including Lyapunov constraints within the MPC problem and show that the stability properties of the linear system (in the observable space) are inherited by the original nonlinear system under certain assumptions. Please note that a rigorous stability analysis of the closed-loop system is out of scope of this work and will be the subject of future research.

### 3. Dynamic model of hydraulic fracturing

Typically, a hydraulic fracturing process consists of two sub-processes: fracture propagation and proppant transport. The corresponding dynamic model is obtained using lubrication theory (fluid momentum) and elasticity equation (rock deformation). We consider one of the classic 2D hydraulic fracturing models, known as the Perkins–Kern–Nordgren (PKN) model [48,49]. This model is usually applied to the cases of long fracture length (hundreds of meters in length, shown as the  $x$ -axis in Fig. 1), constant height ( $y$ -axis) and a small width ( $z$ -axis). The rock bed in which the fractures propagate is assumed to be a homogeneous elastic formation characterized by a constant Young's modulus  $E$  and Poisson's ratio  $\nu$ . Since the fracture length is considered much larger compared to other dimensions, the net pressure is independent of the  $y$ -axis making the cross-section elliptical, the most defining feature of the PKN model.

#### 3.1. Fracture propagation

A brief description of the equations governing the fracture propagation is presented below. The fluid flow rate inside the fracture is determined by the following equation for flow of a Newtonian fluid in an elliptical section using the lubrication theory.

$$\frac{dP}{dx} = -\frac{64\mu q_x}{\pi HW^3} \quad (18)$$

where  $P$  is the net pressure,  $\mu$  is the fracturing fluid viscosity,  $q_x$  is the local flow rate in the horizontal direction  $x$ ,  $H$  is the constant fracture height and  $W$  is the width of the fracture.

The fracture width (the minor axis of the ellipse) caused by the pressure due to the fluid loading is calculated from the solution of the plane strain condition (elasticity equation) as follows

$$W = \frac{2PH(1-\nu^2)}{E} \quad (19)$$

By taking into account the fracture volume changes and the fluid leak-off into the surrounding reservoir, the continuity equation gives the (incompressible) fluid volume conservation as

$$\frac{\partial A}{\partial t} + \frac{\partial q_x}{\partial x} + HU = 0 \quad (20)$$

where  $A = \pi WH/4$  is the cross-sectional area of the elliptic fracture, and  $U(t)$  is the fluid leak-off rate per unit height during the fracture propagation which is given by the Carter's equation [2]

$$U = \frac{2C_{\text{leak}}}{\sqrt{t - \tau(z)}} \quad (21)$$

where  $C_{\text{leak}}$  is the overall leak-off coefficient,  $t$  is the elapsed time since fracturing was initiated, and  $\tau(z)$  is the time at which a specific fracture location gets exposed for the first time.

At the wellbore, the flow rate  $q_z$  is specified, and at the fracture tip  $L(t)$ , the fracture is always closed (i.e., the width of the fracture is zero). These lead to the following two boundary conditions:

$$q_z(0, t) = Q_0; \quad W(L(t), t) = 0, \quad (22)$$

where  $Q_0$  is the fluid injection rate at the wellbore. Initially, the fracture is closed leading to the following initial condition:

$$W(z, 0) = 0 \quad (23)$$

#### 3.2. Proppant transport

To model proppant transport, it is assumed that the injected proppant travels at the carrier fluid's velocity along the horizontal direction governed by advection while simultaneously settling at the fracture bottom forming a proppant bank induced by gravity. The following set of equations is used to describe the proppant transport phenomenon.

The advection of the suspended proppant can be expressed as:

$$\begin{aligned} \frac{\partial(WC)}{\partial t} + \frac{\partial}{\partial x}(WCV_p) &= 0 \\ C(0, t) &= C_0(t) \quad \text{and} \quad C(x, 0) = 0 \end{aligned} \quad (24)$$

where  $C(x, t)$  is the volumetric proppant concentration inside the fracture,  $C_0(t)$  is the injected proppant concentration at the wellbore. The interaction between the individual proppant particles is assumed to be negligible because of the low proppant concentration. The drag and gravitational forces acting on proppant particles are considered for proppant settling. The relationship between the velocity of an individual proppant particle,  $V_p$ , the velocity of the fluid,  $V$ , and gravitational settling velocity,  $V_s$ , is given by [50]:

$$V_p = V - (1 - C)V_s \quad (25)$$

The gravity-induced proppant settling velocity  $V_s$  can be computed as [51]

$$V_s = \frac{(1 - C)^2(\rho_{sd} - \rho_f)gd^2}{10^{1.82C} 18\mu} \quad (26)$$

where  $\rho_{sd}$  is the proppant particle density,  $\rho_f$  is the pure fluid density,  $g$  is the gravitational acceleration constant,  $d$  is the proppant diameter, and  $\mu$  is the fracture fluid viscosity whose relationship with concentration can be modeled through the following empirical expression [52]:

$$\mu(C) = \mu_0 \left( 1 - \frac{C}{C_{\text{max}}} \right)^{-\alpha} \quad (27)$$

where  $\mu_0$  is the pure fluid viscosity,  $\alpha$  is an exponent in the range of 1.2 to 1.8, and  $C_{\text{max}}$  is the maximum theoretical concentration determined by  $C_{\text{max}} = (1 - \phi)\rho_{sd}$  where  $\phi$  is the proppant bank porosity. The evolution of proppant bank height,  $\delta$ , by the settling flux is described by [53],

$$\frac{d(\delta W)}{dt} = \frac{CV_s W}{(1 - \phi)} \quad (28)$$

where there is initially no proppant bank, so the initial condition is that  $\delta(z, 0) = 0$ . Please note that due to dilute suspension, as long as the operation is carried out for a short period of time the proppant bank height will remain much smaller than the fracture height ( $\delta \ll H$ ).

An important characteristic of the hydraulic fracturing process is that as the fracture propagates in the lateral direction, the

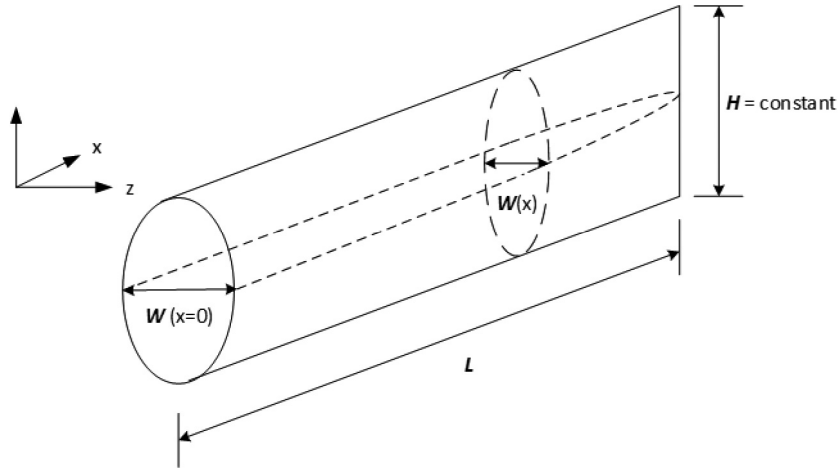


Fig. 1. The PKN fracture model considered in this work.

system boundary changes making the spatial domain of interest time-dependent. An efficient coupling of multiple nonlinear equations that describe the important physical phenomena in hydraulic fracturing systems is essential to update important variables at each time step. Due to the moving boundary nature of the problem, the number of equations to be solved grows as a fracture treatment continues, significantly increasing the computational requirements. Therefore, in this work we explore the use of Koopman operator theory to determine a linear representation of the nonlinear dynamics such that the use of established linear control design methodologies becomes readily applicable.

#### 4. Application to hydraulic fracturing

In practice, the ultimate goal of hydraulic fracturing is to increase the productivity of a stimulated well which is influenced by two important factors: fracture geometry and proppant concentration. Depending on the geological composition of the rock formation, different fracture geometries with the same propped volume may lead to different fracture productivities. For example, in a high-permeability formation, a wide and short fracture is preferred, while in a low-permeability formation, a narrow and long fracture is preferred [3]. Additionally, when the pumping is stopped, the suspended proppant settles and the surrounding rock's natural stress closes the fracture opening thereby trapping the proppant inside. This trapped proppant provides a highly conductive (permeable) channel for effective extraction of oil and gas. Within this regard, here we consider two control problems whose objectives are to achieve (1) desired fracture geometry and (2) uniform proppant concentration at the end of pumping, respectively.

##### 4.1. Regulating fracture geometry

To demonstrate and evaluate the performance of the Koopman system identification method outlined in Section 2.2, we first applied it to regulate the fracture propagation during a hydraulic fracturing process. In the following, we describe in detail the model identification, model performance evaluation and controller design.

###### 4.1.1. Koopman model identification

The first step in the system identification method is to collect and construct the required data matrices. The fracture propagation dynamics can be obtained by solving part of the high-fidelity model described by Eqs. (18)–(23). To deal with the nonlinearity

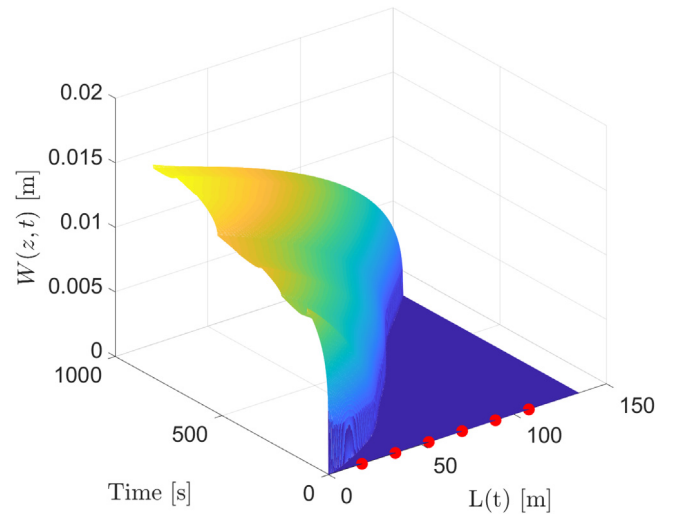


Fig. 2. Spatio-temporal evolution of fracture width. The red markers indicate sensor placement for width measurements. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and moving boundary nature, a novel in-house numerical scheme based on a periodic remeshing strategy was used (we refer the readers to [10] for more details). The values of various process parameters used in our calculations were obtained from [8]:  $H = 20$  m,  $\mu = 0.56$  Pa · s,  $E = 5 \times 10^3$  MPa,  $\nu = 0.2$ , and  $C_{\text{leak}} = 6.3 \times 10^{-5}$  m/s<sup>1/2</sup>. The fracture propagation was terminated at 135 m, and the spatial domain was discretized with each grid point having a size of 0.3 m, resulting in a total of  $n = 451$  points. The spatio-temporal evolution of the fracture width for a sample injection flow rate is shown in Fig. 2. It can be observed that the growth rate of the fracture width is very high in the beginning but it slows down with time. At the same time, the fracture length grows steadily at a constant rate. The moving boundary nature of the problem is evident in the above figure with the zero values at a spatial point indicating that the fracture has not yet propagated to that location.

Because the hydraulic fracturing process takes place deep below the earth's surface, the availability of width measurements at all the spatial locations is not guaranteed. In fact, the only practically available measurements are the fracture length and fracture width at the wellbore determined via the processed

micro-seismic and downhole pressure data, respectively. Therefore, in this work, we assumed only a partially observed system; the average fracture width,  $W_{avg}(t)$ , fracture length,  $L(t)$  and fracture width at six uniformly spaced locations,  $[W_1(t), \dots, W_6(t)]$ , obtained from the numerical experiment were considered as the true process measurements, i.e.,  $x(t) = [W_{avg}(t), L(t), W_1(t), \dots, W_6(t)]^T \in \mathbb{R}^8$ .

The data required to construct the Koopman linear model was collected from 100 simulated trajectories over varying operating time periods with each numerical experiment terminated when the fracture propagates to a total length of 135 m. A total of 1000 (synchronous) time samples per trajectory were used to populate the required data matrices. Each trajectory starts with the same initial condition given in Eq. (23), and was subjected to an input signal generated randomly over each sampling time. The manipulated inputs were bounded as  $\{\underline{u}, \bar{u}\} = \{0.03, 0.06\} \text{ m}^3/\text{s}$ . This results in the data matrices  $X$  and  $Y$  of size  $8 \times 10^5$  and the matrix  $U$  of size  $1 \times 10^5$ . A 50% train–test split was then created to partition the collected data set into training and validation sets.

**Remark 3.** Please note that in the case of a real hydraulic fracturing process, the unmeasurable states such as local fracture width, proppant concentration, etc. can be estimated by using the high-fidelity model and a state observer such as Kalman Filter (see [9, 10] for reference). Alternatively, one can design a state estimator directly for the Koopman linear model using linear observer design methodologies [28,29]. When the high-fidelity model is unknown, one can use classical system identification techniques such as time-delayed measurements to construct the data matrices required for Koopman operator approximation (see [19, 33]).

Once the data has been collected, the next step is to choose a set of basis functions that span the Koopman subspace to which the system dynamics are lifted. In this case, we chose a basis of monomials of the system states with the total degree less than or equal to 2 (all possible linear and bilinear combinations) as follows:

$$\begin{aligned} \phi_i(x) &\in \left\{ \prod_{j=1}^8 x_j^{m_j} \mid (m_1, \dots, m_8) \in \mathbb{Z}_0^+, \sum_{j=1}^8 m_j \leq 2 \right\} \\ \Rightarrow \phi(x) &= [1, W_{avg}, L, W_1, \dots, W_6, W_{avg}^2, W_{avg}L, \dots, W_6^2]^T \end{aligned} \quad (29)$$

where  $\mathbb{Z}_0^+$  denotes the set of non-negative integers. Therefore, the total number of functions in the basis  $\{\phi_i\}_{i=1}^N$  is equal to  $N = 45$ . This choice was motivated by the presence of polynomial terms in the governing equations. One observation here is that when the order of the monomials is 1, they represent the states themselves; in other words, this basis is *state-inclusive*. Consequently, (any of) the original physical states can be recovered by a trivial linear transformation of the lifted states back to the state space as discussed in Section 2.2. For unknown systems, there is no established way to select the dictionary although some canonical choices like radial basis functions have been proposed as a good starting point [20]. As a matter of fact, in the next example we will see that a canonical basis like the polynomial basis selected here fails to accurately capture the dynamics. In such cases, system knowledge becomes paramount.

Using the collected data and the constructed basis  $\phi$ , we determined approximate linear predictors for the fracture propagation model described in Eqs. (18)–(23). We then evaluated its accuracy by comparing the Koopman model predictions to each of the validation data sets, one of which is shown in Fig. 3. From the figure it can be seen that a relatively good agreement between the true model and the identified linear model was achieved with

respect to the test data. To assess this quantitatively, goodness of fit for the trajectory of system states starting from the same initial condition subjected to random input profiles was calculated and averaged over  $N_{total} = 100$  simulations using the average relative root mean squared error (RMSE) defined as:

$$\begin{aligned} \text{RMSE} &= \frac{\|x - \hat{x}\|_{fro}}{\|x\|_{fro}} \\ \text{Avg. RMSE} &= \frac{1}{N_{total}} \sum_{i=1}^{N_{total}} \text{RMSE}_i \end{aligned} \quad (30)$$

where  $\|\cdot\|_{fro}$  is the Frobenius norm. The RMSE details of the validation experiments are presented in Table 1. Based on these results, we can say that the Koopman model consistently captures the real behavior of all eight measured states of the system.

Please note that the developed Koopman models are valid for the fixed values of model parameters presented in the manuscript. When the system parameters change, one has to determine a new model in the new parameter space. For example, in the case of a hydraulic fracturing process, when modeling two different reservoirs, the parameters that are subject to change are the rock properties such as Young's modulus, porosity, permeability, etc. Since the underlying flow physics remain the same, the basis functions will be kept constant, and the Koopman matrices will be computed independently for the reservoirs using the data collected from each reservoir.

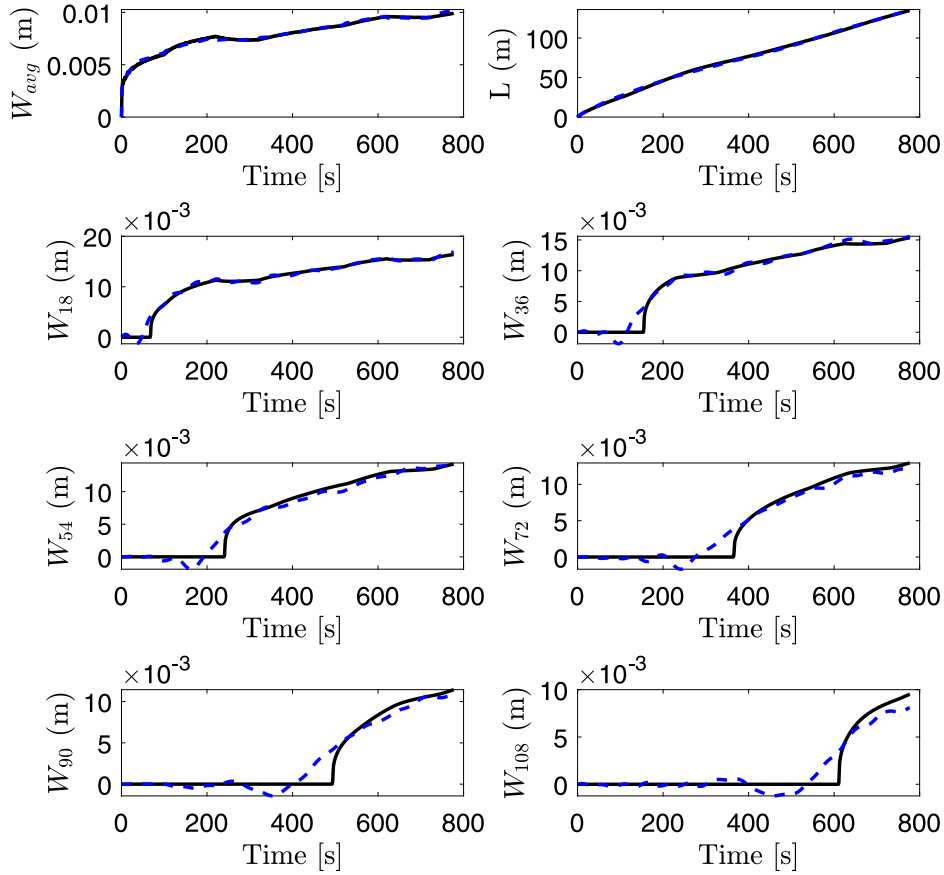
#### 4.1.2. Closed-loop results

Now that an accurate model is identified, the Koopman-based MPC presented in Section 2.3 was formulated to achieve the desired geometry whose objective is to minimize the squared deviation of the fracture length from its set-point at the end of propagation. To prevent early termination of hydraulic fracturing due to tip-screen out, the average fracture width at the end of the operation must be greater than a pre-specified target value; this is considered as the state constraint. So, the MPC controller solves the following optimization problem of the closed-loop operation at sampling time step  $k$ :

$$\begin{aligned} &\text{Minimize}_{u_k, \dots, u_{T-1}} (D\phi_T - L_{target})^2 \\ \text{s.t.} \quad &\phi_{k+j} = A\phi_{k+j-1} + Bu_{k+j-1}, \quad j = 1, \dots, T-k \\ &\phi_k = \phi(x_k) \\ &E\phi_T \leq -W_{target} \\ &\underline{u} \leq u \leq \bar{u} \end{aligned} \quad (31)$$

where  $D \in \mathbb{R}^{1 \times N}$  and  $E \in \mathbb{R}^{1 \times N}$  are matrices that project the lifted state back to the original state space to obtain fracture length and average width, respectively,  $\underline{u}$  and  $\bar{u}$  denote the lower and upper bounds, respectively, on the manipulated input (i.e., the injection flow rate),  $L_{target}$  and  $W_{target}$  are the desired fracture length and average width, respectively,  $k$  is the current time step, and  $T$  denotes the total sampling time steps. The solution of this problem defines a feedback control law  $u = u^*(x_k)$  where only the first value is applied to the closed-loop system for the next sampling time period  $t \in [t_k, t_{k+1})$ , and the procedure is repeated until the end of operation.

To design the above MPC, the optimal fracture geometry which will maximize the productivity of a stimulated well for a given amount of proppant particles was obtained using the UFD scheme. Specifically, the total mass of proppant to be injected was taken as 48,000 kg over the entire treatment. For this specified amount, the corresponding optimal fracture length and width determined by UFD are  $L_{target} = 135$  m and  $W_{target} = 5.4$  mm, respectively. Please refer to [10] for more details on these target values. The control objective is to generate a fracture having a



**Fig. 3.** The measured output of the system (black) superimposed with the predictions of the Koopman-based model (blue dotted) given the same initial condition and control inputs.

**Table 1**  
RMSE values of validation experiments.

| Case                   | Max. RMSE | Min. RMSE | Avg. RMSE |
|------------------------|-----------|-----------|-----------|
| Fracture width         | 6.82      | 0.78      | 3.79      |
| Proppant concentration | 9.77      | 4.77      | 7.19      |

total length equal to  $L_{target}$  while keeping the average fracture width greater than the optimal width  $W_{target}$  at the end of the treatment.

Since the basis was state-inclusive, the matrices  $D$  and  $E$  used in the objective function and state constraints are trivial. Specifically, based on Eq. (29),  $D \in \mathbb{R}^{1 \times N} = [0 \ 1 \ 0_{1 \times N-2}]$  and  $E \in \mathbb{R}^{1 \times N} = [0_{1 \times 2} \ 1 \ 0_{1 \times N-3}]$ , where  $0_{1 \times j}$  is a row matrix whose elements are equal to 0, were used to determine fracture length and average width, respectively. In the closed-loop simulation, the sampling time  $\Delta$  between consecutive measurements was considered to be 100 s and the total process duration to be 800 s; thus, the injection schedule was partitioned into a total of  $T = 8$  stages (i.e.,  $k = 1, \dots, 8$ ) with the duration of each step given by the sampling time,  $\Delta$ . In the beginning of each stage, the state measurements were assumed to be available, which were lifted to the Koopman subspace and used to predict the estimates of the future states via the developed approximate model. The optimization problem was then solved in a shrinking horizon manner (since the control objective was to regulate the output at final time) in this lifted space and the control inputs were computed.

Starting from the initial point, the closed-loop simulation results from  $t = 0$  to  $t = 800$  s are shown in Fig. 4. Using Koopman-based MPC controller, the closed-loop trajectory of the

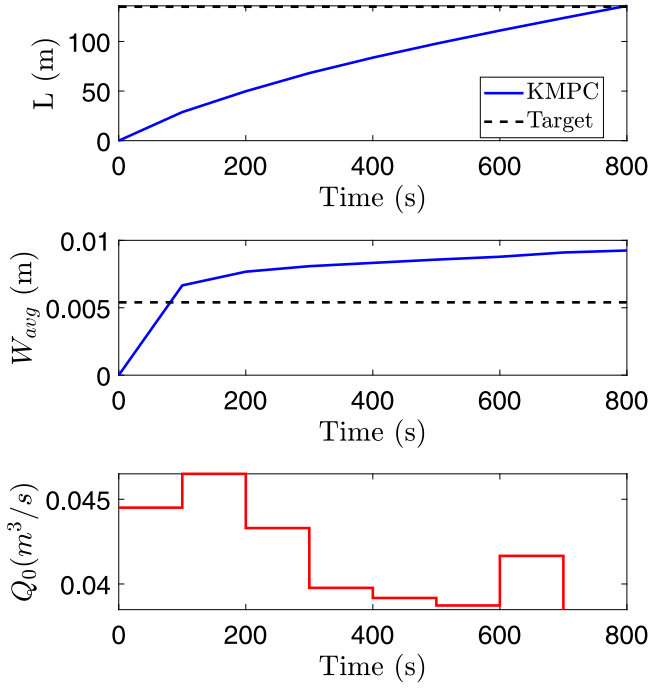
fracture length was able to converge to its set-point at the end of the treatment. Additionally, the average fracture width was able to satisfy the state constraint to prevent tip-screen out.

#### 4.2. Regulating spatial proppant concentration distribution

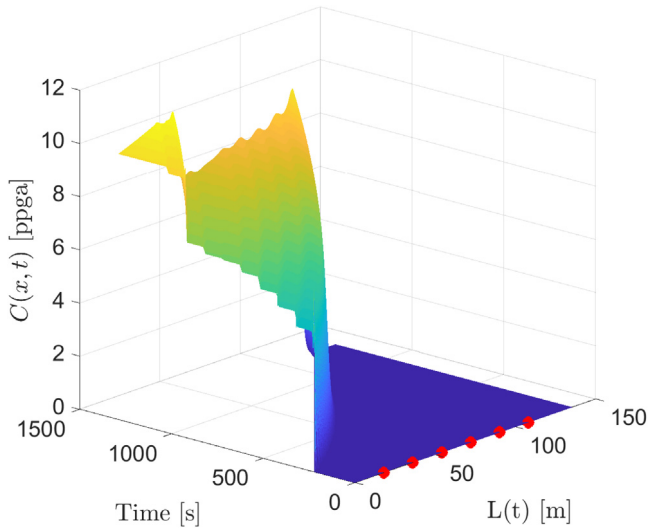
In this section, we consider the design of an optimal pumping schedule by considering the coupled dynamics of fracture propagation as well as proppant transport. The high fidelity model is now described by Eqs. (18)–(28) and is considerably more complex than the case of fracture propagation alone.

##### 4.2.1. Koopman model identification

Similar to the previous example, the numerical experiment was carried out for 100 random input profiles to collect the concentration snapshots and construct the required data matrices for system identification. The full-order solution of the proppant concentration inside the fracture for a sample injected concentration ( $u(t) = C_0(t)$ ) is shown in Fig. 5. It was assumed that the spatial concentration data at the six uniform locations is available as real time measurements, i.e.,  $x \in \mathbb{R}^6$ . The manipulated inputs were bounded as  $\{\underline{u}, \bar{u}\} = \{0, 10\}$  ppga where ppga denotes the pounds of proppant per gallon of fracturing fluid. This results in the data matrices  $X$  and  $Y$  of size  $6 \times 10^5$  and the matrix  $U$  of size



**Fig. 4.** Closed-loop trajectories of fracture length, average width and inputs determined by the MPC.



**Fig. 5.** Spatio-temporal evolution of proppant concentration. The red markers indicate sensor placement for measurements. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$1 \times 10^5$ . A 50% train–test split was then created to partition the collected data set into training and validation sets.

The next step is to choose a dictionary of observable functions to form the Koopman basis. However, due to increased complexity in the dynamics, the choice of basis functions for this case study is not trivial. Several numerical experiments (the results of which are not presented here) revealed that using a canonical basis failed to accurately reproduce the concentration dynamics observed here. Specifically, polynomial basis with an order up to 5 and gaussian radial basis functions have been observed to fail when used as Koopman basis. In such cases, it is really helpful to incorporate any system knowledge to populate the

basis. Specifically, based on the concentration data available from the numerical experiments, it can be observed that the system closely resembles an empirical first-order plus dead time (FOPDT) model. Since the fracture tip moves in the horizontal direction as the process proceeds (moving boundary), there exists a time delay in the proppant concentration injected at the wellbore (input) to reach a particular location. This phenomenon has also been verified in our previous work [54] where sparse regression was used to develop an interpretable reduced-order model for proppant concentration data. The sparse regression-based ROM showed that the proppant concentration at a location inside the fracture is a linear function of its neighboring locations. This can be attributed to simultaneous fracture propagation and fluid leak-off in the hydraulic fracturing process. More specifically, let us suppose that the fracture has propagated to a location  $x$ . As the proppant is transported to that location, its concentration increases with time. The increase in concentration is also contributed by the fluid leaking off into the surrounding reservoir. At the same time, due to simultaneous fracture propagation, a new spatial location, say the location  $x + 1$ , becomes available, which facilitates the transport of the proppant away from the location  $x$ . Therefore, the concentration at the location  $x$  is slightly decreased or remains constant depending on the magnitude of the input. This phenomenon can be clearly seen in the spatial concentration profile shown in Fig. 5. Based on these observations, the Koopman basis was chosen using linear functions of the concentrations at the six locations. Therefore, the total number of functions in the basis  $\{\phi_i\}_{i=1}^N$  is equal to  $N = 6$ . Note that this is again a state-inclusive basis.

$$\phi_i(x) \in \left\{ \prod_{j=1}^6 x_j^{m_j} \mid (m_1, \dots, m_6) \in \mathbb{Z}_0^+, \sum_{j=1}^6 m_j = 1 \right\} \quad (32)$$

$$\Rightarrow \phi(x) = [C_1 \ C_2 \ \dots \ C_6]^T$$

However, the complexity of this system actually lies in the way that the input is handled, not in the states themselves. If a simple linear basis was used without proper handling of the input, the obtained result failed to accurately represent the system. Since it takes different times for the inlet concentration to reach different locations inside the fracture, this time delay in the input variable was incorporated into the model implicitly by constructing 6 auxiliary inputs (i.e., one for each location) based on the time it takes for the inlet concentration to reach the specific location,  $u \in \mathbb{R}^6$ . The numerical values for input time-delays were determined from the simulation data.

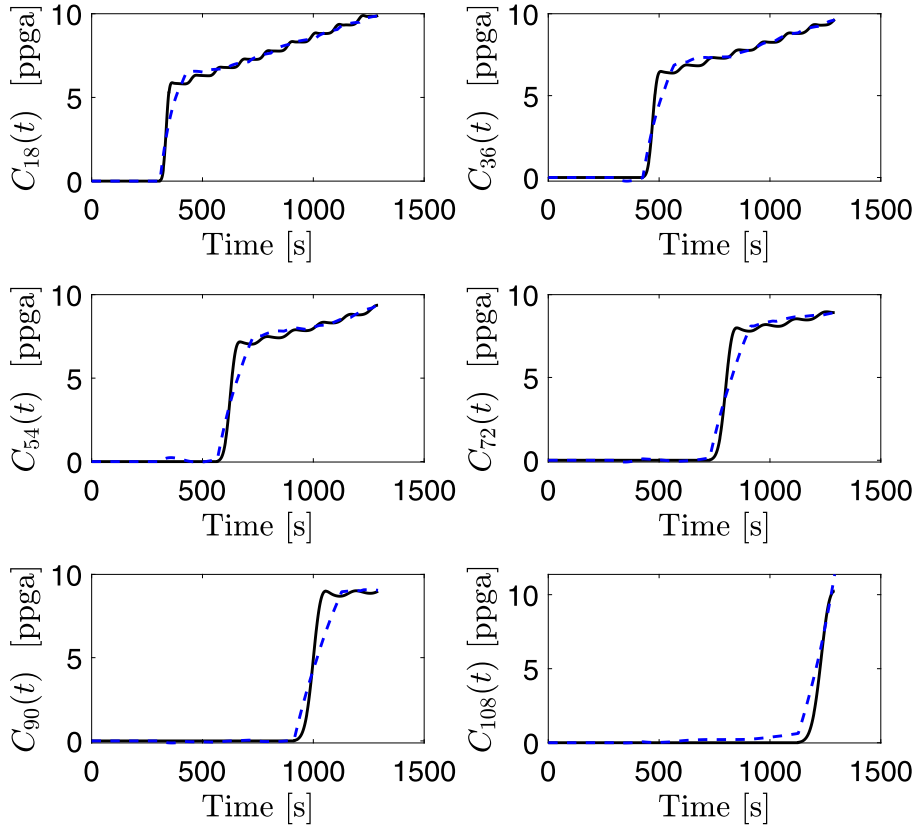
The validation results of the identified Koopman linear model are shown in Fig. 6. It can be seen from the figure that the Koopman model was able to identify the measured concentration dynamics fairly accurately. This was further verified from the low Avg. RMSE value of 7.19% when 100 trajectories were validated with different inputs. Please refer to Table 1 for more details.

#### 4.2.2. Closed-loop results

Based on the optimal length and width computed in the previous example in Section 4.1, a target concentration can be computed as follows:

$$C_{\text{target}} = \frac{M_{\text{prop}}}{HL_{\text{target}} W_{\text{target}}} = 9.5 \text{ ppga} \quad (33)$$

The control objective is to achieve the above concentration  $C_{\text{target}}$  uniformly throughout the fracture. There are several constraints associated with the problem. The input profile should increase monotonically to follow the practical implementation in the field ( $C_{0,k} \leq C_{0,k+1}$ ). A material constraint given by  $2Q_0 \Delta \sum C_{0,k} = M_{\text{prop}}$  is also required to ensure the required amount of proppant is injected into the fracture. All these constraints are introduced



**Fig. 6.** The measured output of the system (black) superimposed with the predictions of the Koopman-based model (blue dotted) given the same initial condition and control inputs.

into the MPC problem as linear constraints in a similar manner as described in Section 2.3. The following optimization problem was then solved over the prediction horizon, and the first step of the solution was applied to the high-fidelity model in a sample-and-hold fashion, and the procedure repeated at every sampling time until the end of treatment.

$$\begin{aligned}
 & \text{Minimize} && (D\phi_T - C_{\text{target}})^T Q (D\phi_T - C_{\text{target}}) \\
 & \text{s.t} && u_k, \dots, u_{T-k} \\
 & && \phi_{k+j} = A\phi_{k+j-1} + Bu_{k+j-1}, \quad j = 1, \dots, T-k \\
 & && \phi_k = \phi(x_k) \\
 & && G_u u \leq 0 \\
 & && G_m u = M_{\text{prop}} \\
 & && \underline{u} \leq u \leq \bar{u}
 \end{aligned} \tag{34}$$

where  $D$  is the identity matrix in this example. The cost matrix  $Q$  was taken to be identity. The input constraint matrix  $G_u \in \mathbb{R}^{T-k \times T-k}$  is an upper bi-diagonal matrix with  $G_u(j, j) = 1$ ,  $G_u(j, j+1) = -1 \forall j = 1, \dots, T-k$ . Essentially, this constraint specifies that the injected concentration at a time step  $k+1$  is greater than or equal to that of  $k$ , i.e., an increasing input profile. This constraint was used to obtain a practical pumping schedule, one that closely resembles the field. The material constraint vector  $G_m$  is given by  $G_m \in \mathbb{R}^{1 \times T} = 2Q_0\Delta * \mathbf{1}_{1 \times T}$  where  $\mathbf{1}_{1 \times j}$  denotes a row vector whose elements are equal to 1. This constraint ensures that the total mass of injected proppant, given by the cumulative input concentration, is equal to the pre-specified value of  $M_{\text{prop}}$ .

The results of the Koopman-based MPC are presented in Fig. 7. From the figure it can be seen that the derived controller was able to regulate the final proppant concentration to the pre-specified target value throughout the fracture. The pumping schedule (i.e., the input concentration profile required to achieve the desired

proppant concentration) for the corresponding process parameters over the entire operation time is shown in Fig. 7b. Please note that the developed approximate models considered only measurements at 6 (out of 451) locations; in the control problem, we do not necessarily require knowledge of the entire system state but only of some observations which are used in the controller. We see that even by using partial observations, the linear model was able to control the PDE fairly accurately. Moreover, the dimension of the optimization problem is reduced because instead of using a high-dimensional discretization and a higher order time integrator, we use a matrix vector product to predict the dynamics of the observed states.

## 5. Conclusions

In this work, we successfully applied the Koopman operator theory for system identification and feedback control of a hydraulic fracturing process. We studied two cases: (1) fracture propagation and (2) proppant transport. The results showed that in both the examples presented, the generated linear models were able to accurately predict the evolution of all the observed states and resulted in a very small Avg. RMSE value (3.79% and 7.19%) over 100 trajectories with different inputs. One major difference in the examples is in the choice of the Koopman basis. In the case of fracture propagation, since the dynamics were relatively simple, a canonical basis like the bilinear basis functions was able to accurately capture the observed dynamics. However, the case of proppant concentration is more sophisticated where the dynamics are highly coupled and nonlinear. In such cases, it is observed that a priori process knowledge can be leveraged to select the required basis of observables. Additionally, sparse regression based dictionary identification can be used as a precursor to obtain the most relevant functions in the basis.

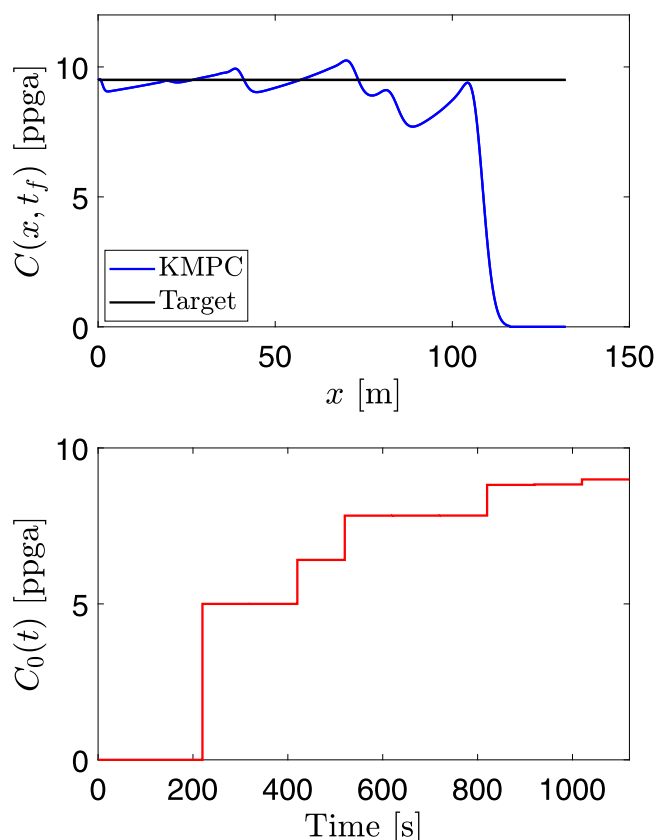


Fig. 7. Final proppant concentration and input trajectories of the closed-loop process determined by the MPC.

The obtained Koopman models for both the cases were then used to design feedback controllers to regulate the fracture geometry and proppant concentration. Due to the superior predictive capabilities of the linear models, the derived controllers achieved their corresponding control objectives; the final desired fracture length while satisfying the constraint on average fracture width and uniform proppant concentration at the end of the fracturing treatment. Moreover, because of its linear structure, the Koopman model resulted in a convex quadratic MPC problem that is amenable to be solved using any of the available linear MPC solvers. Although, future work in this direction will need to certify these controllers, providing guarantees on stability, and a rigorous quantification of error and uncertainty bounds, the current work shows the potential of the operator-theoretic framework for approximation and control of chemical and petroleum engineering processes.

#### CRedit authorship contribution statement

**Abhinav Narasingam:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Visualization. **Joseph Sang-II Kwon:** Conceptualization, Formal analysis, Resources, Writing - review & editing, Supervision, Project administration, Funding acquisition.

#### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgments

The authors gratefully acknowledge financial support from the National Science Foundation, USA (CBET-1804407), the Department of Energy, USA (DE-EE0007888-10-8), the Texas A&M Energy Institute, USA, and the Artie McFerrin Department of Chemical Engineering, Texas A&M University, USA.

#### References

- [1] P. Stevens, The Shale Gas Revolution. Developments and Changes, Chatham House Briefing Paper, 2012.
- [2] M.J. Economides, K.G. Nolte, Reservoir Stimulation, Wiley, Chichester, 2000.
- [3] M.J. Economides, R.E. Oligney, P. Valko, Unified Fracture Design, Orsa Press, 2002.
- [4] S. Bhattacharya, M. Nikolaou, M.J. Economides, Unified fracture design for very low permeability reservoirs, *J. Nat. Gas Sci. Eng.* 9 (2012) 184–195.
- [5] K.G. Nolte, Determination of proppant and fluid schedules from fracturing-pressure decline, *SPE Prod. Eng.* 1 (1986) 255–265.
- [6] H. Gu, J. Desroches, New pump schedule generator for hydraulic fracturing treatment design, in: SPE Latin American and Caribbean Petroleum Engineering Conference, Port-of-Spain, Trinidad and Tobago, 2003, SPE Drilling and Completion.
- [7] E.V. Dontsov, A.P. Peirce, A new technique for proppant schedule design, *Hydraul. Fract. J.* 1 (2014) 1–8.
- [8] Q. Gu, K.A. Hoo, Model-based closed-loop control of the hydraulic fracturing process, *Ind. Eng. Chem. Res.* 54 (5) (2015) 1585–1594.
- [9] P. Siddhamshtetty, S. Liu, P. Valko, J.S. Kwon, Feedback control of proppant bank heights during hydraulic fracturing for enhanced productivity in shale formations, *AIChE J.* 64 (2017) 1638–1650.
- [10] P. Siddhamshtetty, S. Yang, J.S. Kwon, Modeling of hydraulic fracturing and designing of online pumping schedules to achieve uniform proppant concentration in conventional oil reservoirs, *Comput. Chem. Eng.* 114 (2018) 306–317.
- [11] A. Narasingam, P. Siddhamshtetty, J.S. Kwon, Temporal clustering for order reduction of nonlinear parabolic PDE systems with time-dependent spatial domains: Application to a hydraulic fracturing process, *AIChE J.* 63 (9) (2017) 3818–3831.
- [12] A. Narasingam, J.S. Kwon, Development of local dynamic mode decomposition with control: Application to model predictive control of hydraulic fracturing, *Comput. Chem. Eng.* 106 (2017) 501–511.
- [13] A. Narasingam, P. Siddhamshtetty, J.S. Kwon, Handling spatial heterogeneity in reservoir parameters using proper orthogonal decomposition based ensemble Kalman filter for model-based feedback control of hydraulic fracturing, *Ind. Eng. Chem. Res.* 57 (11) (2018) 3977–3989.
- [14] M.S.F. Bangi, A. Narasingam, P. Siddhamshtetty, J.S. Kwon, Enlarging the domain of attraction of local dynamic mode decomposition with control technique: Application to hydraulic fracturing, *Ind. Eng. Chem. Res.* 58 (14) (2019) 5588–5601.
- [15] Y. Lan, I. Mezić, Linearization in the large of nonlinear systems and Koopman operator spectrum, *Physica D* 242 (2013) 42–53.
- [16] B.O. Koopman, Hamiltonian systems and transformation in Hilbert space, *Proc. Natl. Acad. Sci.* 17 (5) (1931) 315–318.
- [17] A. Mauroy, I. Mezić, On the use of Fourier averages to compute the global isochrons of (quasi) periodic dynamics, *Chaos* 22 (3) (2012) 033112.
- [18] P.J. Schmid, Dynamic mode decomposition of numerical and experimental data, *J. Fluid Mech.* 656 (2010) 5–28.
- [19] J.H. Tu, D.M. Luchtenburg, C.W. Rowley, On dynamic mode decomposition: Theory and Applications, *J. Comput. Dyn.* 1 (2014) 391–421.
- [20] M.O. Williams, C.W. Rowley, I.G. Kevrekidis, A data-driven approximation of the Koopman operator: Extending dynamic mode decomposition, *J. Nonlinear Sci.* 25 (6) (2015) 1307–1346.
- [21] I. Mezić, Spectral properties of dynamical systems, model reduction and decompositions, *Nonlinear Dynam.* 41 (1–3) (2005) 309–325.
- [22] M. Grilli, P.J. Schmid, S. Hickel, N.A. Adams, Analysis of unsteady behaviour in shockwave turbulent boundary layer interaction, *J. Fluid Mech.* 700 (2012) 16–28.
- [23] J. Proctor, P. Echhoff, Discovering dynamic patterns from infectious disease data using dynamic mode decomposition, *Int. Health* 7 (2015) 139–145.
- [24] B.W. Brunton, L.A. Johnson, J.G. Ojemann, J.N. Kutz, Extracting spatial-temporal coherent patterns in large-scale neural recordings using dynamic mode decomposition, *J. Neurosci. Methods* 258 (2016) 1–15.
- [25] E. Berger, M. Sastuba, D. Vogt, B. Jung, H.B. Amor, Estimation of perturbations in robotic behavior using dynamic mode decomposition, *J. Adv. Robot.* 29 (5) (2015) 331–343.
- [26] A. Mauroy, I. Mezić, Global stability analysis using the eigenfunctions of the Koopman operator, *IEEE Trans. Automat. Control* 61 (11) (2016) 3356–3369.

- [27] A. Mauroy, I. Mezić, J. Moehlis, Isostables, isochrons, and Koopman spectrum for the action-angle representation of stable fixed point dynamics, *Physica D* 261 (2013) 19–30.
- [28] A. Surana, Koopman operator based observer synthesis for control-affine nonlinear systems, in: 55th IEEE Conf. Decision and Control, CDC, Las Vegas, USA, 2016, pp. 6492–6499.
- [29] A. Surana, A. Banaszuk, Linear observer synthesis for nonlinear systems using Koopman operator framework, *IFAC-PapersOnLine* 49 (18) (2016) 716–723.
- [30] A. Surana, M.O. Williams, M. Morari, A. Banaszuk, Koopman operator framework for constrained state estimation, in: 56th IEEE Conf. Decision and Control, CDC, Melbourne, Australia, 2017, pp. 94–101.
- [31] E. Kaiser, J.N. Kutz, S.L. Brunton, Data-driven discovery of Koopman eigenfunctions for control, 2017, [arXiv:1707.01146](https://arxiv.org/abs/1707.01146).
- [32] H. Arbabi, M. Korda, I. Mezić, A data-driven koopman model predictive control framework for nonlinear flows, 2018, [arXiv:1804.05291](https://arxiv.org/abs/1804.05291).
- [33] H. Arbabi, M. Korda, I. Mezić, A data-driven koopman model predictive control framework for nonlinear partial differential equations, in: 57th IEEE Conf. Decision and Control, CDC, Miami Beach, USA, 2018, pp. 6409–6414.
- [34] M. Korda, I. Mezić, Linear predictors for nonlinear dynamical systems: Koopman operator meets model predictive control, *Automatica* 93 (2018) 149–160.
- [35] M. Korda, Y. Susuki, I. Mezić, Power grid transient stabilization using Koopman model predictive control, 2018, [arXiv:1803.10744](https://arxiv.org/abs/1803.10744).
- [36] S. Hanke, S. Peitz, O. Wallscheid, S. Klus, J. Böcker, M. Dellnitz, Koopman operator based finite-set model predictive control for electrical drives, 2018, [arXiv:1804.00854](https://arxiv.org/abs/1804.00854).
- [37] A. Sootla, A. Mauroy, D. Ernst, Optimal control formulation of pulse-based control using Koopman operator, *Automatica* 91 (2018) 217–224.
- [38] S. Peitz, S. Klus, Koopman operator-based model reduction for switched-system control of PDEs, *Automatica* 106 (2019) 184–191.
- [39] B. Huang, X. Ma, M. Vaidya, Feedback stabilization using Koopman operator, in: 57th IEEE Conf. Decision and Control, CDC, Miami Beach, USA, 2018, pp. 6434–6439.
- [40] A. Narasingam, J.S. Kwon, Koopman Lyapunov-based model predictive control of nonlinear chemical process systems, *AIChE J.* 65 (11) (2019) e16743.
- [41] R. Mohr, I. Mezić, Construction of eigenfunctions for scalar-type operators via Laplace averages with connections to Koopman operator, 2014, [arXiv:1403.6559](https://arxiv.org/abs/1403.6559).
- [42] M.O. Williams, M.S. Hemati, S.T. Dawson, I.G. Kevrekidis, Extending data-driven Koopman analysis to actuated systems, *IFAC-PapersOnLine* 49 (8) (2016) 704–709.
- [43] J.L. Proctor, S.L. Brunton, J.N. Kutz, Generalizing Koopman theory to allow for inputs and control, *SIAM J. Appl. Dyn. Syst.* 17 (1) (2018) 909–930.
- [44] M. Korda, I. Mezić, On convergence of extended dynamic mode decomposition to the Koopman operator, *J. Nonlinear Sci.* 28 (2) (2018) 687–710.
- [45] M. Ellis, H. Durand, P. Christofides, A tutorial review of economic model predictive control methods, *J. Process Control* 24 (8) (2014) 1156–1178.
- [46] S. Klus, P. Koltai, C. Schütte, On the numerical approximation of the Perron–Frobenius and Koopman operator, *J. Comput. Dyn.* 3 (1) (2016) 51–79.
- [47] Y. Chen, U. Vaidya, Sample complexity for nonlinear stochastic dynamics, in: 2019 American Control Conference, ACC, 2019, pp. 3526–3531.
- [48] T.K. Perkins, L.R. Kern, Widths of hydraulic fractures, *J. Pet. Technol.* 13 (1961) 937–949.
- [49] R. Nordgren, Propagation of a vertical hydraulic fracture, *Soc. Petrol. Eng. J.* 12 (1972) 306–314.
- [50] J. Adachi, E. Siebrits, A. Pierce, J. Desroches, Computer simulation of hydraulic fractures, *Int. J. Rock Mech. Min. Sci.* 44 (2007) 739–757.
- [51] A. Daneshy, Numerical solution of sand transport in hydraulic fracturing, *J. Pet. Technol.* 30 (1978) 132–140.
- [52] R. Barree, M. Conway, Experimental and numerical modeling of convective proppant transport, *J. Pet. Technol.* 47 (1995) 216–222.
- [53] E.J. Novotny, Proppant transport, in: Proceedings of the 52nd SPE Annual Technical Conference and Exhibition, SPE 6813, Denver, CO, 1977.
- [54] A. Narasingam, J.S. Kwon, Data-driven identification of interpretable reduced-order models using sparse regression, *Comput. Chem. Eng.* 119 (2018) 101–111.