Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnline CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 53-2 (2020) 8814–8819

Robust PID Control of Multicompartment Lung Mechanics Model Using Runge-Kutta Neural Disturbance Observer ⋆

Erdem Dilmen *

* *Mechatronics Engineering Department, Pamukkale University, Denizli, Turkey (e-mail: edilmen@pau.edu.tr)*

Abstract: This paper proposes Runge-Kutta neural disturbance observer to enhance the robustness of PID control of a system with general multicompartment lung mechanics. It is designed to observe the states of a particular type continous time, single-input single-output system where the states cannot be measured but can be observed through the single output and there exists parametric uncertainity or disturbance affecting the underlying system. It utilizes artificial neural network to estimate the disturbance online. Once an accurate disturbance estimation is obtained, it is incorporated in the system state equation and passed through the well-known Runge-Kutta integrator to predict the state values. Hence, the predicted states are obtained considering the disturbance and more robust state observation is achieved. The proposed observer is simple and easy to implement. Adaptation of the neural network is performed using gradient descent with an adaptive learning rate which guarantees convergence. The simulation results demonstrate that the proposed observer gains a significant success in enhancing the robustness of PID control at even high level of disturbance. Note that, multicompartment lung mechanics system is a stand-in model that can mimic the behavior of human lung. Thus, it is appropriate for hardware-in-the-loop simulation which opens a path to the real-patient-tests of mechanical respiratory systems in the future.

Keywords: Multicompartment lung mechanics, PID, artificial neural network, disturbance observer, robust control, Runge-Kutta discretization

1. INTRODUCTION

Human lungs may sometimes contract a serious illness. Such phenomenon may cause malfunctioning of the lungs, leading to a respiratory failure. It is the result of incapability of the human respiratory system delivering an adequate gas excahnge of carbon dioxide (CO_2) and oxygen through the network of capillaries alongside the alveoli. At that point, mechanical ventilation systems arise as a solution.

Operation of the mechanical ventilators can be basically classified in three groups such as volume-controlled, pressure-controlled and dual-controlled (Tobin, 2006). The primary goal of mechanical respiratory systems is, by applying a limited input pressure, to maintain adequate minute ventilation which is the tidal volume multiplied by number of breaths per minute (West and Luks, 2015). In volume-controlled ventilation, both the tidal volume and number of breaths are determined by the clinician who is looking after the patient. Hence, the ventilator operates in accordance to those pre-determined ventilation parameters. However, in pressure-controlled ventilation, the tidal volume is not directly controlled. Instead, the ventilator determines the adequate input pressure that will be able to inflate the lung and make it reach the level of desired tidal volume. Note that, the tidal volume is dependent on the pressure applied and the compliance of the lung. Thus, minute ventilation is dependent on the compliance parmeters of the lung as well. If the compliance is subject to change, then, change in the tidal volume, and depending on that, change in the minute ventilation are inevitable. Thus, robust control algorithms which takes into account the parameter uncertainities are necessary in the mechanical ventilation research.

In this paper, we propose a novel disturbance observation approach for a particular type continuous time, single-input single-ouput (SISO) system where the states are observable through the output. It exploits the Runge-Kutta (RK) discretization method for obtaining a discretized version of the original system. Observer dynamics are derived using the system nominal model in combination with a standard feedforward artificial neural network (ANN). At each sampling time, we estimate the unknown disturbance term affecting the underlying system. We use it within the observer dynamics which is embedded into the RK formulation. Thus, we observe the system states considering the effect of disturbance by utilizing machine learning and RK integration method in the state observation scheme. The proposed observer structure is simple and easy to implement in both computers and embedded hardware. We employ it within PID control to demonstrate how the proposed observer can enhance the robustness of PID at even high levels of disturbance. Also, online training of ANN is performed using gradient descent with an adaptive learning rate which guarantees the convergence of ANN

⋆ This paper is supported by Pamukkale University Scientific Reseearch Projects Council (BAP).

10.1016/j.ifacol.2020.12.1390

parameters. Consequently, convergence of the observer is maintained as well.

Organization of the paper is as follows. In Sect. 2, problem definition and fourth order RK discretization method will be given. Sect. 3 briefly introduces ANN as the artificial model employed for disturbance estimation in this paper. Sect. 4 covers the proposed approach of robust state observation in details. Also in that section, convergence of the ANN parameters by using an adaptive learning rate is presented. Sect. 5 presents the simulation results. Finally, the paper is concluded in Sect. 6 where also the future works are declared.

2. PROBLEM DEFINITION AND RUNGE-KUTTA DISCRETIZATION

This section briefly explains the appraoch of how the robust state observation task is treated. For this, definition of the problem and fourth order Runge-Kutta discretization are given.

2.1 Problem Definition

Consider that we have a continous time, SISO system that is subject to parametric uncertainities and disturbance. In this case, its dynamics are governed by the combination of nominal system model and the terms denoting the parametric uncertainities and external disturbance respectively. State space model for such a system is given below.

$$\begin{aligned}\frac{d\mathbf{x}(t)}{dt} =&\mathbf{f}(\mathbf{x}(t)) + \bar{\mathbf{f}}(\mathbf{x}(t))+\\ &(\mathbf{g}(\mathbf{x}(t)) + \bar{\mathbf{g}}(\mathbf{x}(t)))u(t) + \mathbf{d}_{ext}(t)\\ y(t) =&h(\mathbf{x}(t))\end{aligned} \tag{1}$$

In (1), $\mathbf{x} \in \Re^{d_x}$ denotes the system state vector given the initial value $\mathbf{x}_0 = \mathbf{x}(0)$ while $u \in \Re$ denotes the input. $y \in \Re$ denotes the system output and $h(\cdot)$ is the system function associated with output. Note the different notations used for system functions in the state equation. $\mathbf{f}(\cdot) : \Re^{d_x} \to \Re^{d_x}$ and $\mathbf{g}(\cdot) : \Re^{d_x} \to \Re^{d_x}$ represent the system functions associated with the states and input in the nominal model of the system. On the other hand, $\bar{\mathbf{f}}(\cdot) : \Re^{d_x} \to \Re^{d_x}$ and $\bar{\mathbf{g}}(\cdot) : \Re^{d_x} \to \Re^{d_x}$ denote the uncertainity in the system model (1). Last, $\mathbf{d}_{ext} \in \Re^{d_x}$ denotes the external disturbance on the system.

If we treat the parametric uncertainity as internal disturbance, $\mathbf{d}_{int} \in \Re^{d_x}$, and combine that internal disturbance with the external one, we can rewrite (1) as follows.

$$\begin{aligned}\frac{d\mathbf{x}(t)}{dt} &= \mathbf{f}(\mathbf{x}(t)) + \mathbf{g}(\mathbf{x}(t))u(t) + \mathbf{d}(t)\\ \mathbf{d}(t) &= \mathbf{d}_{int}(t) + \mathbf{d}_{ext}(t)\\ \mathbf{d}_{int}(t) &= \bar{\mathbf{f}}(\mathbf{x}(t)) + \bar{\mathbf{g}}(\mathbf{x}(t))u(t)\\ y(t) &= h(\mathbf{x}(t))\end{aligned} \tag{2}$$

In (2), we collect the internal and external disturbances within one disturbance term $\mathbf{d} \in \Re^{d_x}$. Consequently, we write the overall system model as a combination of the nominal model and the total disturbance $\mathbf{d}$. We assume that we have measurements of the system output y but we cannot measure the states individually. Hence, to construct an appropriate control architecture, we need to observe the states through the output. It is assumed that each state x_i, $i = 1, \ldots, d_x$, can be observed via the measured output y. We can develop an accurate state observer once we have an accurate estimate of $\mathbf{d}$. To develop a state observer, we first need to obtain a descretized version of the originally continuous time system (2).

2.2 Runge-Kutta Discretization

RK method enables us to descretize the continous time system (2). With a sampling period T_s, the continous time output y is sampled at the time instants nT_s where $n = \{0, 1, 2, \ldots\} \in \mathbb{Z}$ (Iplikci, 2013). We can obtain discrete samples of the input similarly as well. Let us show the discrete time variables $\mathbf{x}[n] = \mathbf{x}(nT_s)$ and $u[n] = u(nT_s)$ for the state vector and input respectively. Using the fourth order RK method, we can write the discretized system equations as follows.

$$\begin{aligned}&\mathbf{x}[n+1] = \mathbf{x}[n] + \boldsymbol{\delta}_x\big(\mathbf{x}[n], u[n+1], \mathbf{d}[n+1]\big)\\ &\boldsymbol{\delta}_x\big(\mathbf{x}[n], u[n+1], \mathbf{d}[n+1]\big) = \frac{1}{6}K_1 + \frac{1}{3}K_2 + \frac{1}{3}K_3 + \frac{1}{6}K_4\\ &y[n+1] = h\big(\mathbf{x}[n+1]\big)\end{aligned} \tag{3}$$

where

$$\begin{aligned}K_1 &= T_s\big(\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u + \mathbf{d}[n+1]\big)_{\mathbf{x}=\mathbf{x}[n], u=u[n+1]}\\ K_2 &= T_s\big(\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u + \mathbf{d}[n+1]\big)_{\mathbf{x}=\mathbf{x}[n]+\frac{1}{2}K_1, u=u[n+1]}\\ K_3 &= T_s\big(\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u + \mathbf{d}[n+1]\big)_{\mathbf{x}=\mathbf{x}[n]+\frac{1}{2}K_2, u=u[n+1]}\\ K_4 &= T_s\big(\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u + \mathbf{d}[n+1]\big)_{\mathbf{x}=\mathbf{x}[n]+K_3, u=u[n+1]}\end{aligned} \tag{4}$$

The discretized system of equations (3) and (4) will be utilized in Sect. 4 to obtain an accurate estimate of the total disturbance term $\mathbf{d}$ by using ANN. Hence, a robust observer for the system states $\mathbf{x}$ could be developed taking into account the effect of total disturbance.

3. ARTIFICIAL NEURAL NETWORK

This section briefly introduces a standard feedforward ANN which is adopted for estimating the total disturbance term $\mathbf{d}$ in the discretized system given by (3) and (4). ANN is one of the well-known function approximators that is structured by artificial neurons. Mathematically, ANN model output in response to a feature vector $\mathbf{h}_n \in \Re^{d_h}$, corresponding to the n_{th} sampling time, $n = \{0, 1, 2, \ldots\} \in \mathbb{Z}$, can be written as

$$\hat{\mathbf{d}}_n = \mathbf{w}_o\boldsymbol{\phi}(\mathbf{w}_h\mathbf{h}_n + \mathbf{b}_h) + b_o \tag{5}$$

where $\mathbf{w}_o \in \Re^{d_x \times s}$ and $\mathbf{w}_h \in \Re^{s \times d_h}$ are the output layer and hidden layer weight matrices, respectively. Note that, s denotes the number of neurons while d_h denotes dimension of the feature vector. Also, T denotes the transpose operator for the vectors/matrices. In (5), $\mathbf{b}_h \in \Re^s$ and $b_o \in \Re$ are the vector of hidden layer biases and output bias respectively. Note that, the actual input vector to ANN is the feature vector $\mathbf{h}$. Feature vector holds the meaningful data to be presented to the ANN. The details for that will be given in Sect. 4.2.

Remember that, we want to estimate the disturbance term using ANN. Thus, output of the ANN is denoted by the approximated disturbance, $\hat{\mathbf{d}} \approx \mathbf{d}$, in (5). In (5), $\boldsymbol{\phi}(.)$

denotes the activation function which fires the neurons. A Gaussian function,

$$\boldsymbol{\phi}(\mathbf{a}) = e^{-0.5((\mathbf{a}-\boldsymbol{\mu})./\boldsymbol{\sigma}).^2} \tag{6}$$

is used as the activation function in this paper. In (6), $\boldsymbol{\mu} \in \Re^s$ and $\boldsymbol{\sigma} \in \Re^s$ denote the center and width parameters of the Gaussian activation function respectively for any arbitrary input $\mathbf{a} \in \Re^s$. Since, exponential function is computing with vectors, the notation './' and '.^{2}' stands for the elementwise division and elementwise square operators for the vectors. A single hidden layer, multi-input multi-output (MIMO) ANN topology is shown in Fig. 1 as a general representation. Last, we construct a parameter

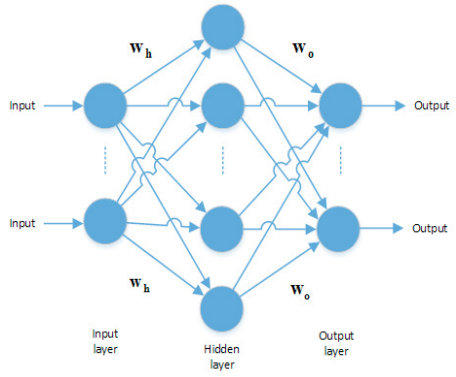

Fig. 1. Feedforward ANN model with a single hidden layer.

vector $\boldsymbol{\theta} \in \Re^{d_\theta}$ which holds all the model parameters coming from the weight matrices and bias vectors $\mathbf{w}_h$, $\mathbf{b}_h$, $\mathbf{w}_o$, b_o, $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ where $d_\theta = s \times (d_h + d_x + 3) + 1$.

4. PROPOSED RUNGE-KUTTA NEURAL DISTURBANCE OBERVER

This section presents the proposed approach of developing a robust state observer which enables us to observe the unmeasured states of an observable, continous time system which can be represented by (2).

4.1 Proposed Observer Dynamics

What we propose is referred to as a disturbance observer due to two reasons. First, it estimates the total disturbance $\mathbf{d}$ on the system by utilizing the measured output data y and the nominal model of the system. Output prediction error is backpropagated to the artificial model which is employed for estimating the total disturbance term in the RK-discretized system given by (3) and (4). Note that, in this study, the chosen artificial model is ANN. Secondly, the proposed disturbance observer incorporates the estimated disturbance in the observer dynamics to yield a robust state observation. In other words, it actually estimates the states $\mathbf{x}$ and the total disturbance term $\mathbf{d}$ simultaneously. For the rest of the paper, when we say 'disturbance' we will actually mean the 'total disturbance'. Let us denote the estimated disturbance by $\hat{\mathbf{d}}$ and the estimated states by $\hat{\mathbf{x}}$. Joint state and disturbance estimation by the proposed RK neural (RKN) disturbance observer can be written for the $(n+1)_{th}$ sampling time in three stages as follows. Note that, the function $\boldsymbol{\delta}_x(\cdot,\cdot,\cdot)$ associated with RK discretization is adopted from (3) and (4).

Error generation:

$$\begin{aligned}
\hat{\mathbf{d}}^-[n+1] &= \mathbf{w}_o\boldsymbol{\phi}(\mathbf{w}_h\mathbf{h}_{n+1} + \mathbf{b}_h) + b_o \\
\hat{\mathbf{x}}^-[n+1] &= \hat{\mathbf{x}}[n] + \boldsymbol{\delta}_x\big(\hat{\mathbf{x}}[n], u[n+1], \hat{\mathbf{d}}^-[n+1]\big) \\
\hat{y}^-[n+1] &= h\big(\hat{\mathbf{x}}^-[n+1]\big) \\
e^-[n+1] &= y[n+1] - \hat{y}^-[n+1]
\end{aligned} \tag{7}$$

ANN parameter update:

$$\boldsymbol{\theta}_{n+1} = \boldsymbol{\theta}_n + \Delta\boldsymbol{\theta}_{n+1} \tag{8}$$

Joint state and disturbance estimation:

$$\begin{aligned}
\hat{\mathbf{d}}[n+1] &= \mathbf{w}_o\boldsymbol{\phi}(\mathbf{w}_h\mathbf{h}_{n+1} + \mathbf{b}_h) + b_o \\
\hat{\mathbf{x}}[n+1] &= \hat{\mathbf{x}}[n] + \boldsymbol{\delta}_x\big(\hat{\mathbf{x}}[n], u[n+1], \hat{\mathbf{d}}[n+1]\big) \\
\hat{y}[n+1] &= h\big(\hat{\mathbf{x}}[n+1]\big)
\end{aligned} \tag{9}$$

The term $\hat{\mathbf{d}}^-$ in (7) can be interpreted as an a-priori disturbance estimate just created to obtain an output prediction error which is necessary to update the parameters of ANN, and thus, to obtain an accurate estimate of the disturbance, denoted by $\hat{\mathbf{d}}$ in (9). In computation of both $\hat{\mathbf{d}}^-$ and $\hat{\mathbf{d}}$, the feature vector $\mathbf{h}_{n+1}$ which is used as input within the ANN output equation is constructed as expressed in Sect. 4.2. Also note that, the ANN parameters used for computing $\hat{\mathbf{d}}^-$ in (7) belongs to the ANN parameter vector $\boldsymbol{\theta}_n$ which was defined in Sect. 3 and contains the parameters which were last updated at the n_{th} sampling time. After the parameter update (8) associated with the sampling time $n+1$, the updated parameters of ANN, which are contained in the vector $\boldsymbol{\theta}_{n+1}$, are used to compute the accurate estimate of disturbance, denoted by $\hat{\mathbf{d}}_{n+1}$ in (9). The respective parameter update term $\Delta\boldsymbol{\theta}_{n+1}$ is obtained by backpropagation of the output prediction error $e^-[n+1]$ in (7) and given in Sect. 4.3.

4.2 Feature Vector Construction

Since we estimate the disturbance at each sampling time, we actually desire to capture the trend in the disturbance over time. We can achieve this goal using the measured input-output history which is mostly the case in time series prediction studies. Thus, we construct a nonlinear autoregressive with exogeneous input (NARX) data model to construct an approriate feature vector $\mathbf{h}$ to ANN whose output equation is given by (5). Here, we make an intuitive choice such that, instead of using the output history, we prefer using the historical data of system states. Since they are unmeasurable, we will use the history of estimated states by the proposed RKN disturbance observer within the NARX data. The adopted feature vector in this paper associated with the $(n+1)_{th}$ sampling time is constructed as follows: $\mathbf{h}_{n+1} = \big[u[n+1], \ldots, u[n+1-n_p], \hat{x}_1[n], \ldots, \hat{x}_1[n+1-n_p], \ldots, \hat{x}_{d_x}[n], \ldots, \hat{x}_{d_x}[n+1-n_p]\big]^T \in \Re^{d_h}$ where $d_h = n_p \times (d_x + 1) + 1$. Note that, $\hat{x}_i$, $i = 1, \ldots, d_x$, denote each individual state estimated by the RKN disturbance observer while n_p denotes the past horizon which is also referred to as the NARX data order.

4.3 RK-Derived Equations of Backpropagation

We define a quadratic loss function of the output prediction error at the $(n+1)_{th}$ sampling time.

$$\begin{aligned} \min_{\boldsymbol{\theta}} L_{n+1} &= \frac{1}{2} e^{-}[n+1]^2 \\ e^{-}[n+1] &= y[n+1] - \hat{y}^{-}[n+1] \end{aligned} \tag{10}$$

We will use gradient descent method with an adaptive learning rate for the update of ANN parameters. Hence, $\Delta\boldsymbol{\theta}_{n+1} = -\alpha \frac{\partial L_{n+1}}{\partial \boldsymbol{\theta}_n}$ will be used where $\alpha > 0$ is the adaptive learning rate and will be detailed in Sect. 4.4. Let us write the gradient expression.

$$\begin{aligned} \frac{\partial L_{n+1}}{\partial \boldsymbol{\theta}_n} &= \frac{\partial L_{n+1}}{\partial e^{-}[n+1]} \frac{\partial e^{-}[n+1]}{\partial \boldsymbol{\theta}_n} = e^{-}[n+1] \frac{\partial e^{-}[n+1]}{\partial \boldsymbol{\theta}_n} \\ &= e^{-}[n+1] \frac{\partial e^{-}[n+1]}{\partial \hat{\mathbf{x}}^{-}[n+1]} \frac{\partial \hat{\mathbf{x}}^{-}[n+1]}{\partial \hat{\mathbf{d}}^{-}[n+1]} \frac{\partial \hat{\mathbf{d}}^{-}[n+1]}{\partial \boldsymbol{\theta}_n} \end{aligned} \tag{11}$$

In (11), $\frac{\partial e^{-}[n+1]}{\partial \hat{\mathbf{x}}^{-}[n+1]} = -\frac{\partial h\left(\hat{\mathbf{x}}^{-}[n+1]\right)}{\partial \hat{\mathbf{x}}^{-}[n+1]}$. Also, the term $\frac{\partial \hat{\mathbf{d}}^{-}[n+1]}{\partial \boldsymbol{\theta}_n}$ is actually partial dervative of ANN output wrt its parameters since the term $\hat{\mathbf{d}}^{-}[n+1]$ is computed by ANN in (7). It is trivial to obtain that partial derivative and due to the limited space, its explicit expression is omitted. Other than that, the partial derivative term $\frac{\partial \hat{\mathbf{x}}^{-}[n+1]}{\partial \hat{\mathbf{d}}^{-}[n+1]}$ is computed utilizing the RK-discretized system formulation by using (2), (3), (4) and (7) as follows. Let us show the time indices as subscripts, e.g., $\hat{x}^{-}[n+1]$ as $\hat{x}^{-}_{n+1}$, $\hat{\mathbf{d}}^{-}[n+1]$ as $\hat{\mathbf{d}}^{-}_{n+1}$ and so on, for simpler notation.

$$\frac{\partial \hat{x}^{-}_{n+1}}{\partial \hat{\mathbf{d}}^{-}_{n+1}} = \frac{1}{6}\frac{\partial K_1}{\partial \hat{\mathbf{d}}^{-}_{n+1}} + \frac{1}{3}\frac{\partial K_2}{\partial \hat{\mathbf{d}}^{-}_{n+1}} + \frac{1}{3}\frac{\partial K_3}{\partial \hat{\mathbf{d}}^{-}_{n+1}} + \frac{1}{6}\frac{\partial K_4}{\partial \hat{\mathbf{d}}^{-}_{n+1}} \tag{12}$$

where

$$\begin{aligned} \frac{\partial K_1}{\partial \hat{\mathbf{d}}^{-}_{n+1}} &= T_s \frac{\partial \dot{\mathbf{x}}}{\partial \hat{\mathbf{d}}^{-}_{n+1}}, \quad \frac{\partial \hat{\mathbf{d}}^{-}_{n+1}}{\partial \mathbf{x}} = \left.\frac{\partial \hat{\mathbf{d}}^{-}_{n+1}}{\partial \mathbf{x}}\right|_{\mathbf{x}=\hat{\mathbf{x}}_n, u=u_{n+1}} \\ \frac{\partial K_2}{\partial \hat{\mathbf{d}}^{-}_{n+1}} &= T_s \left(0.5 \left.\frac{\partial \dot{\mathbf{x}}}{\partial \mathbf{x}}\right|_{K_2} \frac{\partial K_1}{\partial \hat{\mathbf{d}}^{-}_{n+1}} + \frac{\partial \dot{\mathbf{x}}}{\partial \hat{\mathbf{d}}^{-}_{n+1}} \right) \\ \left.\frac{\partial \dot{\mathbf{x}}}{\partial \mathbf{x}}\right|_{K_2} &= \frac{\partial \hat{\mathbf{d}}^{-}_{n+1}}{\partial \mathbf{x}} + \left(\frac{\partial \mathbf{f}(\mathbf{x})}{\partial \mathbf{x}} + \frac{\partial \mathbf{g}(\mathbf{x})}{\partial \mathbf{x}} \right)_{\mathbf{x}=\hat{\mathbf{x}}_n + \frac{1}{2}K_1, u=u_{n+1}} \\ \frac{\partial K_3}{\partial \hat{\mathbf{d}}^{-}_{n+1}} &= T_s \left(0.5 \left.\frac{\partial \dot{\mathbf{x}}}{\partial \mathbf{x}}\right|_{K_3} \frac{\partial K_2}{\partial \hat{\mathbf{d}}^{-}_{n+1}} + \frac{\partial \dot{\mathbf{x}}}{\partial \hat{\mathbf{d}}^{-}_{n+1}} \right) \\ \left.\frac{\partial \dot{\mathbf{x}}}{\partial \mathbf{x}}\right|_{K_3} &= \frac{\partial \hat{\mathbf{d}}^{-}_{n+1}}{\partial \mathbf{x}} + \left(\frac{\partial \mathbf{f}(\mathbf{x})}{\partial \mathbf{x}} + \frac{\partial \mathbf{g}(\mathbf{x})}{\partial \mathbf{x}} \right)_{\mathbf{x}=\hat{\mathbf{x}}_n + \frac{1}{2}K_2, u=u_{n+1}} \\ \frac{\partial K_4}{\partial \hat{\mathbf{d}}^{-}_{n+1}} &= T_s \left(\left.\frac{\partial \dot{\mathbf{x}}}{\partial \mathbf{x}}\right|_{K_4} \frac{\partial K_3}{\partial \hat{\mathbf{d}}^{-}_{n+1}} + \frac{\partial \dot{\mathbf{x}}}{\partial \hat{\mathbf{d}}^{-}_{n+1}} \right) \\ \left.\frac{\partial \dot{\mathbf{x}}}{\partial \mathbf{x}}\right|_{K_4} &= \frac{\partial \hat{\mathbf{d}}^{-}_{n+1}}{\partial \mathbf{x}} + \left(\frac{\partial \mathbf{f}(\mathbf{x})}{\partial \mathbf{x}} + \frac{\partial \mathbf{g}(\mathbf{x})}{\partial \mathbf{x}} \right)_{\mathbf{x}=\hat{\mathbf{x}}_n + K_3, u=u_{n+1}} \end{aligned} \tag{13}$$

In (12) and (13), the term $\frac{\partial \dot{\mathbf{x}}}{\partial \mathbf{x}}$ is actually partial derivative of the state equation (2) for the continuous time system wrt to the state vector. Based on that, the partial derivative $\frac{\partial \dot{\mathbf{x}}}{\partial \hat{\mathbf{d}}^{-}_{n+1}} = \mathbf{I} \in \Re^{d_x \times d_x}$ where $\mathbf{I}$ denotes the identitiy matrix. It should be noticed that, the term $\frac{\partial \hat{\mathbf{d}}^{-}_{n+1}}{\partial \mathbf{x}}$ is actually the partial derivative of ANN output wrt to the state vector since $\hat{\mathbf{d}}^{-}_{n+1}$ is computed by ANN in (7). The feature vector $\mathbf{h}_{n+1}$ involves the most recent estimated states $\hat{\mathbf{x}}_n$ due to the definition of feature vector construction in Sect. 4.2. Note that, $\frac{\partial \hat{\mathbf{d}}^{-}_{n+1}}{\partial \mathbf{x}}$ is computed only at the first step of the fourth order RK discretization scheme and its fixed value is used in the following steps through two to four as seen in (13). We can expand that term as

$$\frac{\partial \hat{\mathbf{d}}^{-}_{n+1}}{\partial \mathbf{x}} = \left.\frac{\partial \hat{\mathbf{d}}^{-}_{n+1}}{\partial \mathbf{x}}\right|_{\mathbf{x}=\hat{\mathbf{x}}_n, u=u_{n+1}} = \left.\frac{\partial \hat{\mathbf{d}}^{-}_{n+1}}{\partial \mathbf{h}_{n+1}} \frac{\partial \mathbf{h}_{n+1}}{\partial \hat{\mathbf{x}}_n}\right|_{u=u_{n+1}} \tag{14}$$

Obtaining the explicit expression for (14) is trivial and omitted due to the limited space.

4.4 Convergence of ANN Parameters

Let us adopt the quadratic, positive definite loss function in (10) as our Lyapunov candidate function. For simple notation, we will denote the sampling instants as a subscript. Then, let us write a difference equation for the Lyapunov function considering (11) and the parameter update term $\Delta\boldsymbol{\theta}_{n+1} = -\alpha \frac{\partial L_{n+1}}{\partial \boldsymbol{\theta}_n}$ as follows.

$$\begin{aligned} \Delta L_{n+1} &= L_{n+1} - L_n = \frac{1}{2}\left((e^{-}_{n+1})^2 - (e^{-}_n)^2 \right) \\ &= \Delta e^{-}_{n+1} \left(e^{-}_{n+1} + \frac{1}{2}\Delta e^{-}_{n+1} \right) \end{aligned} \tag{15}$$

where

$$\begin{aligned} \Delta e^{-}_{n+1} &\approxeq \underbrace{\left(\frac{\partial e^{-}_{n+1}}{\partial \boldsymbol{\theta}_n} \right)^T}_{\mathbf{J}^T_{n+1}} \Delta\boldsymbol{\theta}^T_{n+1} \approxeq -\alpha \mathbf{J}^T_{n+1} \frac{\partial L_{n+1}}{\partial \boldsymbol{\theta}_n} \\ &\approxeq -\alpha e^{-}_{n+1} \mathbf{J}^T_{n+1} \mathbf{J}_{n+1} \end{aligned} \tag{16}$$

Substituing Δe^{-}_{n+1} of (16) in (15), and after some algebraic operations, we obtain the condition below which maintains $\Delta L_{n+1} < 0$ and thus, guarantees convergence of the ANN parameters.

$$0 < \alpha < \frac{2}{\mathbf{J}^T_{n+1}\mathbf{J}_{n+1}} \tag{17}$$

Considering the condition (17), in this study, without loss of generality we use the apative learning rate below

$$\alpha = \frac{1}{\mathbf{J}^T_{n+1}\mathbf{J}_{n+1} + \epsilon} \tag{18}$$

where $\mathbf{J}_{n+1}$ is computed at each time instant $n = 0, 1, \ldots$, and so depending on that, the learning rate α is adaptive. The term ϵ is a small positive number (e.g., 0.01) that prevents zero-denominator situation. Hence, it is necessary for numerical stability. Convergence of the ANN parameters provide convergence of the estimated states to real values in finite time.

5. SIMULATION RESULTS

In this study, we adopt a linear model (Hou et al., 2014) of multicompartment lung mechanics as the nominal

model in purpose. It involves the lung compliances and air resistances as constant system parameters. However, during control, we assume that the actual system has time-varying compliances. In addition, we apply an artifical external disturbance $\mathbf{d}_{ext}$ to the underlying system as well. We test our proposed observer in the PID control framework to assess its contribution to robustness.

5.1 Nominal System Model

Let us briefly explain the multicompartment lung mechanics model. It has a dichotomoy that is inspired by human lung where at each generation of a new airway, the airway opens to a subsequent two-branch airway, and so on. If we have a system of γ generations, then we have 2^γ lung compartments. A lung mechanics model for $\gamma = 2$ is given in Fig. 2. In Fig. 2, $R^{in}_{j,i}$ denotes the airway resistance for

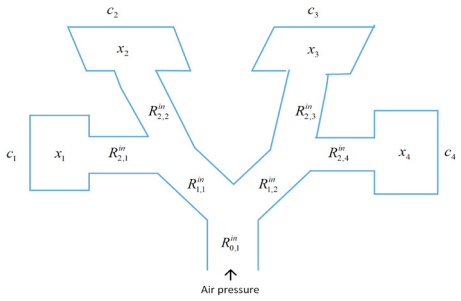

Fig. 2. Lung mechanics model with 4 compartments, $\gamma = 2$.

the i_{th} airway of the j_{th} generation for $j = 0, \dots, \gamma$ and $i = 1, \dots, 2^j$. Note that, the superscript 'in' stands for the inhalation period and its exhalation counterpart is denoted by $R^{ex}_{j,i}$. Also, x_i and c_i for $i = 1, \dots, 2^\gamma$, denote the air volume and the associated compliance of the compartment i. Let us write the state space equations for the model. Note that, the system has switched dynamics since the parameters associated with the inhalation and exhalation periods need not be the same.

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t) \\ y(t) &= \mathbf{1}^T\mathbf{x}(t) \end{aligned} \tag{19}$$

where $\mathbf{1} \in \Re^{d_x}$ is a vector of ones, $\mathbf{x} = [x_1 \dots x_{d_x}]^T \in \Re^{d_x}$ is the state vector, $d_x = 2^\gamma$, $\mathbf{x}(0) = \mathbf{x}_0$ and

$$\begin{cases} \mathbf{A} = -\mathbf{R}_{in}^{-1}\mathbf{C}, & \mathbf{B} = \mathbf{R}_{in}^{-1}\mathbf{1}, & 0 \le t \le T_{in} \\ \mathbf{A} = -\mathbf{R}_{ex}^{-1}\mathbf{C}, & \mathbf{B} = \mathbf{R}_{ex}^{-1}\mathbf{1}, & T_{in} < t \le T_{in} + T_{ex} \end{cases} \tag{20}$$

T_{in} denotes the inhalation period while T_{ex} denotes the exhalation period. Thus, a breathing period is computed as $T = T_{in} + T_{ex}$. System (20) is periodic with the period T and the output $y(t)$ is actually sum of the states, which means, we measure the total lung volume as the system output. Input $u(t)$ to the system is the air pressure applied at the initial airway whose resistance is denoted by $R^{in}_{0,1}$. $\mathbf{C} = diag\,[1/c_1 \dots 1/c_{d_x}]$. Note that, $\mathbf{R}_{in} = \sum_{j=0}^{\gamma}\sum_{i=1}^{2^j} R^{in}_{j,i}\mathbf{Z}_{j,i}\mathbf{Z}^T_{j,i}$ and $\mathbf{R}_{ex} = \sum_{j=0}^{\gamma}\sum_{i=1}^{2^j} R^{ex}_{j,i}\mathbf{Z}_{j,i}\mathbf{Z}^T_{j,i}$ where the l_{th} element of $\mathbf{Z}_{j,i} \in \Re^{d_x}$ is 1 for all $l = (i-1)2^{\gamma-j+1}, (i-1)2^{\gamma-j+2}, \dots, i2^{\gamma-j}$ and 0 otherwise. Last, let us give the parameters for nominal model of the system. We should indicate that we gave a general representation of the respective system by Fig. 2. However, we use a two-compartment lung mechanics model in the simulations ($\gamma = 1$). $R^{in}_{0,1} = 9$ cm H_2O/l/s, $R^{in}_{1,1} = 16$ cm H_2O/l/s and $R^{in}_{1,2} = 16$ cm H_2O/l/s. $R^{ex}_{j,i}$ counterpart is twice the $R^{in}_{j,i}$, $j = 0, 1$ and $i = 1, 2$. Lung compartment compliance nominal values are $c_i = 0.11$/cm H_2O, $i = 1, 2$.

5.2 Illustrative Example

Initial compartment air volumes are set $\mathbf{x}_0 = [0.5, 0]^T$ and initial observer states are $\hat{\mathbf{x}}_0 = [0, 0]^T$ liters. Also, input air pressure has an upper bound $u_{max} = 19$ cm H_2O. Inhalation and exhalation periods are $T_{in} = 2$ s and $T_{ex} = 3$ s respectively while $T_s = 0.1$ s. ANN parameters are initially set to 1e-3. Also, $s = 3$ and $n_p = 3$ are found to bu suitable experimentally. In the simulation, in contrast to the nominal model, the lung compliance parameters are time-varying and denoted by c^{tv}_i, $i = 1, 2$. A time-varying compliance by (21) has a profile which varies between half and full value of the nominal compliance within any of the inhalation or exhalation periods.

$$\begin{aligned} &c^{tv}_i(t) = c_i + \bar{c}_i(t) \\ &\bar{c}_i(t) = -0.5c_i + 0.5c_i sin(2\pi f t') \\ &\begin{cases} t' = mod(t, T), f = 0.5/T_{in} & , mod(t, T) \le T_{in} \\ t' = mod(t, T) - T_{in}, f = 0.5/T_{ex} & , mod(t, T) > T_{in} \end{cases} \end{aligned} \tag{21}$$

The term $\bar{c}_i$ in (21) corresponds to a hard parametric uncertainty which is created artificially. It is used for computation of the internal disturbance term $\mathbf{d}_{int}$ defined in (2). In the simulation, in addition to the internal disturbance $\mathbf{d}_{int}$, an external disturbance $\mathbf{d}_{ext}$ given by (22) is applied to the underlying system as well.

$$\mathrm{d}_{i,ext}(t) = d_o + \sum_{k=1}^{3} d_{m_k} sin\big(2\pi(1/T_{d_k})t\big), i = 1, 2 \tag{22}$$

where $d_o = -0.2$, $T_d = [2, 5, 10]$ s, $d_m = [0.03, 0.02, 0.01]$ and $\mathbf{d}_{ext} = [\,\mathrm{d}_{1,ext},\, \mathrm{d}_{2,ext}]^T$. The input signal produced by the PID controller is formulated as follows where $K_p = 202.34$, $K_i = 1.03$ and $K_d = 2.54$ are found to be suitable by a grid search.

$$\begin{aligned} &u(t) = K_p e(t) + K_d \dot{e}(t) + K_i \int_0^t e(\tau)d\tau \\ &e(t) = y_{ref}(t) - y(t), \dot{e}(t) = \dot{y}_{ref}(t) - \mathbf{1}^T\dot{\hat{\mathbf{x}}}(t), \mathbf{1} \in \Re^{d_x} \end{aligned} \tag{23}$$

Note that, the crucial term in (23) is $\dot{e}(t)$ since it involves $\dot{y}(t) = \mathbf{1}^T\dot{\hat{\mathbf{x}}}(t)$. It is computed using the estimated state vector $\hat{\mathbf{x}}$ in the system state equation given by (2). y_{ref} denotes the reference respiratory pattern. Fig. 3 shows the tracking result. y denotes the system output obtained when PID control based on the proposed RKN disturbance observer is performed, which means, $\dot{e}$ in (23) is computed substituting the estimated states $\hat{\mathbf{x}}$ and disturbance $\hat{\mathbf{d}}$ with $\mathbf{x}$ and $\mathbf{d}$ in (2). On the other hand, y_{nom} denotes the system output for the case where the proposed observer is not active, which means, $\dot{e}(t)$ is computed in a similar way but only the nominal model is considered and disturbance is not taken into account. Note that, when RKN disturbance observer is not used, since the system states are not measurable, we prefer employing a standard Luenberger oberver which is well-known from the literature. Due to

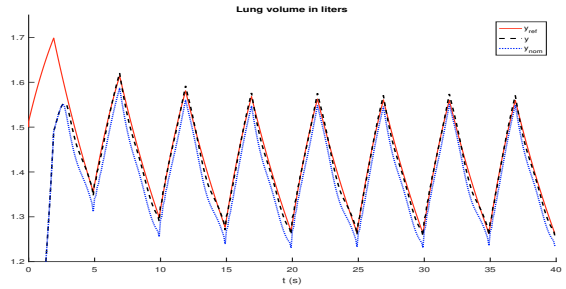

Fig. 3. Tracking result.

the limited space, its details are omitted (Wang et al., 2016). The root mean squarred tracking error values for the cases where observer is active and not are 0.1468 and 0.1553 respectively. It is clear from y_{nom} in Fig. 3 that, controlling a system which is under both internal (parametric uncertainity) and external disturbances by considering its nominal model is not enough to achieve an acceptable tracking performance when the internal and external disturbances are relatively at high level given by (21) and (22). Fortunately, even in such hard case, the proposed disturbance observer can contribute to robustness of the control method employed, resulting in an acceptable control performance. Fig. 4 shows the produced control input signals where u_{nom} is the one produced by considering the nominal model of the system while u is produced when control is perfomed based on the RKN disturbance observer. Fig. 5 shows us that, the proposed

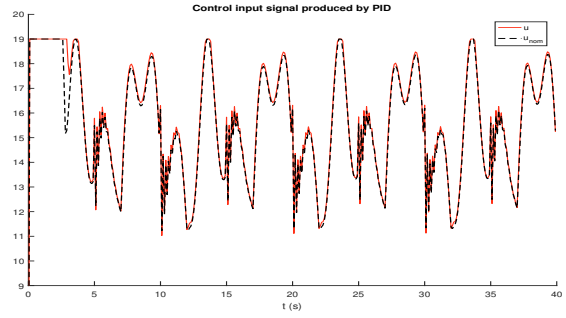

Fig. 4. Produced control input signal.

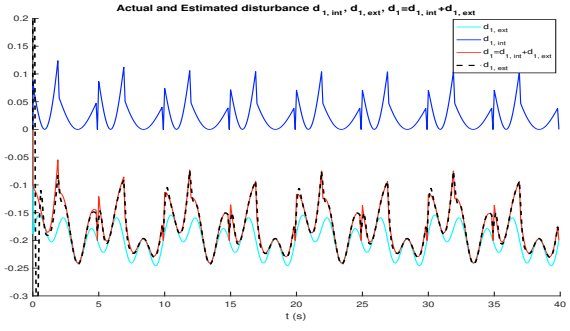

Fig. 5. Disturbance estimation d_1.

RKN disturbance observer can estimate the disturbance $\mathbf{d} = \mathbf{d}_{int} + \mathbf{d}_{ext}$ with an excellent performance and thus, state estimate $\hat{\mathbf{x}}$ converges to the system states $\mathbf{x}$ in a finite time, which leads to a state estimation error $\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}}$ converging to zero, as seen from Fig. 6. Note that in Fig. 5,

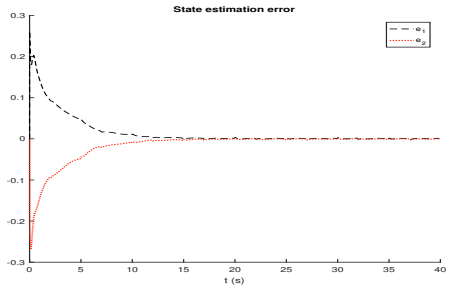

Fig. 6. State estimation error.

only the disturbance d_1 is plotted since we have limited space and disturbance d_2 is similar.

It should be noticed that, output y_{nom} in Fig. 3 tends to approach the reference y_{ref} over time, which is possible by the aid of the integrator term in PID. This may make one think as if y_{nom} could track the reference better when K_i is incerased more. However, we tested several bigger values of K_i and when it gets much bigger, it deteriorates the tracking result.

6. CONCLUSION

In this paper, we propose RKN disturbance observer which estimates the disturbance and system states simultaneously. It both takes into account parametric uncertainity and external disturbance. Hence, it provides robust state observation. It contributes to robustness of the adopted control strategy. It is applied within PID control of a multicompartment lung mechanics model. It manages to capture well the modelling uncertainities in addition to disturbance estimation. Simulation results demonstrate that it can be a useful tool in robust control even at high levels of parametric uncertainity and disturbance. In the future, we plan to employ it within sliding mode control of multicompartment lung mechanics in real time. Also, it will be characterized with noise-filtering capability utilizing the main principles of Kalman filtering technique for noisy conditions.

REFERENCES

Hou, S.P., Meskin, N., and Haddad, W.M. (2014). A general multicompartment lung mechanics model with nonlinear resistance and compliance respiratory parameters. In *2014 American Control Conference*, 566–571.

Iplikci, S. (2013). Runge–Kutta model-based adaptive predictive control mechanism for non-linear processes. *Transactions of the Institute of Measurement and Control*, 35(2), 166–180.

Tobin, M. (2006). Principles and practice of mechanical ventilation, 2nd edition. *Shock*, 26, 426.

Wang, J., Liu, Y., and Sun, C. (2016). Luenberger observer design for state estimation of a linear parabolic distributed parameter system with discrete measurement sensors. In *2016 12th World Congress on Intelligent Control and Automation (WCICA)*, 1123–1128.

West, J.B. and Luks, A.M. (2015). *West's Respiratory Physiology: The Essentials*. LWW, 10th edition.