Identification and Modeling of Distillation Columns From Transient Response Data

Diana Ugryumova*, Gerd Vandersteen
Vrije Universiteit Brussel
dept. ELEC, Pleinlaan 2, 1050 Brussels, Belgium

Bart Huyck[‡§], Filip Logist[‡], Jan Van Impe[‡], Bart De Moor[§]
Katholieke Universiteit Leuven
dept. CIT, W. de Croylaan 46, 3001 Heverlee, Belgium[‡]
dept. ESAT, Kasteelpark Arenberg 10, 3001 Heverlee, Belgium[§]

***Abstract*—Mathematical models that predict the behavior of the column are highly useful to improve the working conditions of distillation columns, i.e. increase throughput or lower energy consumption. They are a prerequisite for advanced control strategies. White-box models of distillation columns already exist, but they are often too time consuming to calibrate in real industrial cases. This paper models a binary distillation column using a black-box model in the frequency domain. One of the identification challenges of distillation columns is the long transient response that causes leakage errors in the frequency domain. This problem is solved using a recently developed local polynomial method. Furthermore, we propose a simple expedient to model the variations in the column response that are caused by the changes in ambient temperature. The results are illustrated using both simulation data of a white-box model and using real measurements.**

***Index Terms*—Identification, frequency response, time-varying systems, transient response.**

I. Introduction

Distillation is the most used separation technique worldwide and is very energy consuming. Hence, it is important to reduce energy consumption as much as possible.

An efficient way of reducing energy consumption in distillation columns is by using accurate models. This helps to predict and control the distillation processes, and therefore helps to reduce the energy consumption.

Several so-called white-box models of distillation columns exist, which make use of balance equations for material and energy, flow dynamics, etc. [1]. These models are in practice often cost and labor intensive to calibrate. The goal here is to find an accurate but simple black-box model of a distillation column to be used, for example, in model predictive control.

There are several challenges in the identification of a distillation column. These include: (a) large time constants, (b) variations due to ambient temperature changes, (c) constraints on the operation and excitation of the column, (d) measurement noise, (e) nonlinearities, (f) time-delays in the response of the column, and (g) the ill-conditioning of the system [2], [3].

(a) A long transient response (large time constant) in the time domain translates into leakage in the frequency domain. Because of leakage errors, the frequency response function (FRF) estimate is biased when using frequency-domain system identification methods [4].

* Corresponding author, Email: diana.ugryumova@vub.ac.be

(b) The varying ambient temperature causes the system's operating point to change. Thus, we cannot use the same model for different ambient temperatures.

(c) Because of safety reasons, there are certain constraints on the operation of the column. Also, several limitations of the column's excitation exist. For example, the duty of the reboiler has to be above a certain minimum for the separation of liquids to take place. At the same time, the reboiler power may not exceed a certain level if the less volatile component has to remain at the bottom of the column.

(d) Measurement noise is always present due to imperfect measurement instruments. Here, we assume that there is noise present on both the input and the output measurements (errors-in-variables framework [4]).

(e) A high-purity distillation column is believed to be strongly nonlinear system. This should be taken into the consideration when selecting a model structure for the distillation column. The pilot-scale distillation column we use is not a high-purity column, allowing us to use a linear model to estimate the column.

(f) Time-delays are present in a distillation column, for example, due to its recycle loops. The time-delays of the pilot-scale column used are in the order of several minutes.

(g) The ill-conditioning of the system means that it has strong directionality [2], [3], i.e. the condition number of the gain matrix is (too) large. In other words, this means that the processes inside a distillation column are strongly interactive. Although this is believed to be a major problem in distillation column identification, we will not address the problem here.

In this paper, we address several of the above mentioned identification problems for distillation columns ((a), (b) and (d) to (f)). Namely, we eliminate the leakage error and estimate the level of nonlinearities and noise using nonparametric identification methods. In addition, we also find a parametric low-order rational form model in the Laplace domain with a time-delay that depends on the ambient temperature. To solve the leakage problem and to estimate the nonparametric FRF together with its error bounds, we use a local polynomial method (LPM). The choice of the system's excitation also plays an important role. Here, we use an optimized random-phase multisine excitation and make sure that it meets the constraints of the column. The LPM eliminates the transient by first fitting a polynomial in the frequency domain to the nonexcited frequencies followed by a transient removal step.

12-139

Ugryumova D., Vandersteen G., Huyck B., Logist F., Van Impe J., De Moor B., ``Identification and Modeling of Distillation Columns From Transient Response Data '', in *Proc. of the IEEE International Instrumentation and Measurement Technology Conference (I²MTC 2012)*, Graz, Austria, May 2012, pp. 2098 - 2103.

Then, using the nonparametric estimate of the FRF, we find a low-order rational form model with time-delay for the system's FRF. Finally, a simple expedient on the reboiler power is used that makes the estimation of a model with parameters depending on the ambient temperature easier. Thus, the goal is to estimate a meta-model of distillation column whose parameters depend on the ambient temperature.

The paper is structured as follows: first, the identification methods to a model distillation columns are explained in Section II. Subsection II-A defines a multisine excitation, while the LPM is explained in Subsection II-B. Parametric modeling is mentioned in Subsection II-C. How to handle temperature-dependence of our system is explained in Subsection II-D. In Section III, a simulation of a white-box model of a pilot-scale binary distillation column, originally developed by [5] and [6], is used to demonstrate the method. Finally, we present the results using real measurement data from a pilot-scale distillation column at KUL-BioTeC in Section IV.

II. Solution method

A distillation column is considered to be a multiple-input-multiple-output (MIMO) nonlinear system whose system dynamics can be affected by the changes in ambient temperature. In this research, we consider an uninsulated pilot-scale distillation column which is not in steady-state; we therefore have to deal with leakage in the frequency domain.

In this section, we explain the identification and modeling methods we use to eliminate the transient, model the varying dynamics due to the ambient temperature, estimate the nonparametric frequency response matrix (FRM) and give an indication of the level of nonlinearities.

Optimized random-phase multisine inputs are used together with the LPM to obtain better nonparametric estimates of the FRM and its variance. This nonparametric FRM is used to fit a low-order rational form parametric model. To account for the changes in ambient temperature, we propose that the parameters of the model vary as a function of the difference between the ambient temperature and the total heat loss of the column.

A. Multisine excitation

Orthogonal random-phase multisine inputs are used to optimally excite the column [7]. Multisines are periodic signals, which in our case have the same amplitude a_i, the same fundamental frequency f_0, the same sample period T_s of one second and zero mean random phase $\phi_{k,i,r}$ for each excited frequency k, input u_i and realization r:

$$u_{i,r}(t) = \sum_{k=1}^{F} a_i cos(2\pi f_0 k t T_s + \phi_{k,i,r}) \qquad (1)$$

Orthogonal random-phase multisines are several realizations of multisines where the matrices of phases for each realization are unitary matrices. The advantage of these orthogonal multisines is their guaranteed optimal conditioning of the input matrix that is used to compute the nonparametric FRM [7].

B. Local polynomial method

A short description of a recently developed nonparametric identification technique, the so-called 'robust' LPM [8], is given in this subsection. Robust LPM is used in combination with an orthogonal multisine excitation to preprocess the data. The LPM helps us to remove the transient, estimates the nonparametric FRM and the variances of FRM that are due to noise and nonlinearities. The LPM starts with the knowledge that the spectra satisfy:

$$U(k) = U_0(k) + N_U(k) + T_{G_U}(\Omega_k),$$
$$Y(k) = G(\Omega_k)U_0(k) + Y_s(k) + N_Y(k) + T_{G_Y}(\Omega_k)$$

where U and Y are the Fourier transforms of the input and output sequences respectively, Ω_k is the generalized frequency at the k-th frequency line, $G(\Omega)$ is the $n_y \times n_u$ FRM, N_U and N_Y are the noise terms, Y_s is the zero mean stochastic nonlinear distortion, T_{G_U} and T_{G_Y} are the transient terms of the system. In this case, $G(\Omega)$ is the best linear approximation of the system, in mean square sense [9].

The robust LP method is used to locally approximate the transient terms and remove them [8]. This comes from the knowledge that the transient in frequency domain is a smooth function of the frequency. Thus, we can approximate the transient $T(\Omega)$ by a polynomial of the frequency, as in:

$$T(\Omega_{k+r}) = T(\Omega_k) + t_1(k)r + t_2(k)r^2 + t_3(k)r^3 + ...$$

We use several periods of orthogonal random-phase multisine signals to excite the system. Transforming this to the frequency domain, we know exactly which frequency lines are excited $Y(k_{\text{ex}})$ and which are nonexcited $Y(k_{\text{nonex}})$. It is reasonable to assume that only the excited frequency lines contain the information about frequency response G and the nonlinearities Y_s. The transient and the noise contributions are present on all the lines, such that:

$$Y(k_{\text{ex}}) = G(\Omega_k)U_0(k) + Y_s(k) + N_Y(k) + T_{G_Y}(\Omega_k)$$
$$Y(k_{\text{nonex}}) = N_Y(k) + T_{G,N}(\Omega_k)$$

In this way, the nonexcited lines are used to locally fit a polynomial to the transient terms and estimate the noise covariance. Next, the estimated transient term is subtracted from all the excited lines, leaving:

$$Y(k_{\text{ex}}) = G(\Omega_k)U_0(k) + Y_s(k) + N_Y(k)$$

The robust LPM uses n_u experiments for each realization. At least two realizations are needed to differentiate between Y_s and N_Y. As a result, different realizations are used to calculate the average FRM and the total covariance on the FRM estimate, $Y_s + N_Y$. For more details about LPM, see [8], [9].

C. Parametric model

Eventually, our goal is to obtain a parametric model of the distillation column. Thus, using the nonparametric FRM estimate and its variance, found with LPM, we make a parametric model of our system using MATLAB function

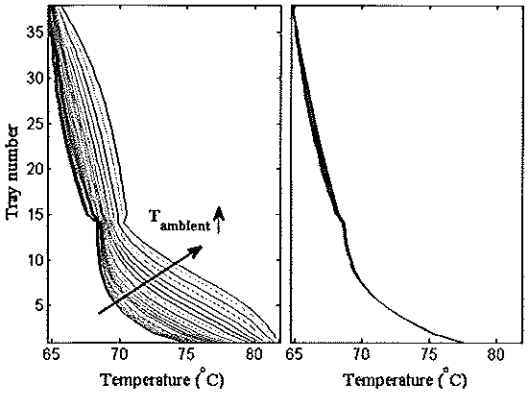


Fig. 1. Plot of the temperature inside the column as a function of the tray number; to the left: operation of the column varies with ambient temperature; to the right: the reboiler power is adjusted to the variations in the ambient temperature such that the operational temperatures of the column become more or less constant.

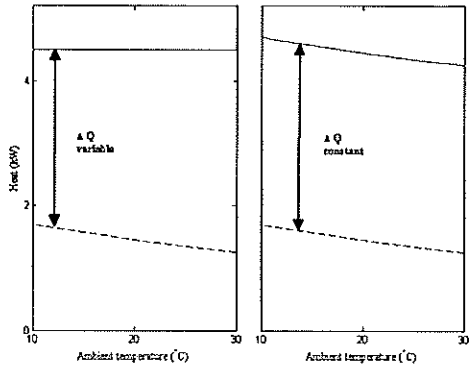


Fig. 2. Plot of the reboiler power (solid line) vs. heat loss over the column (dashed line); ΔQ denotes the difference between the power in the reboiler and the heat loss over the column. Left: ΔQ is changing with ambient temperature; right: heat losses are compensated in the reboiler power to make ΔQ constant with respect to the variations in the ambient temperature.

elis from the *fdident* package [10]. This model describes the distillation column by its poles and zeros and a time-delay term.

Initially, these parametric models are obtained for a constant ambient temperature. By varying the ambient temperature, the operating point of the system changes. Thus, we estimate a parametric model for several operating points of the system.

D. Temperature varying system

Distillation columns are most often not isolated from their environment. Because of this, the amount of heat lost over the whole column during the operation is directly related to the ambient temperature. This relation is shown in the following equation:

$$Q_{\text{loss}}(T_{\text{ambient}}) = \sum_{i=0}^{\#\text{Trays}} S_i \cdot C_{\text{HT}} \cdot (T_{\text{tray}_i} - T_{\text{ambient}}) \quad (2)$$

where S_i is the outer surface of the i^{th} part of the column, C_{HT} is a coefficient of the heat transfer between the distillation column and the ambient environment, T_{tray_i} is the temperature on tray i, where $\text{tray}_{i=0}$ is the reboiler, and T_{ambient} is the environmental temperature. Hence, the ambient temperature has an impact on the operation of the distillation column. This impact of the ambient temperature can be significant for small pilot-scale columns with a large surface-to-volume ratio, whereas the impact is less pronounced for large industrial columns with a small surface-to-volume ratio. To enable an enhanced control, models can incorporate changes due to the ambient temperature.

Initially, the idea for a model structure that accounts for the ambient temperature was a model with parameters and delay term varying with ambient temperature. We denote such a model as $\mathcal{M}(Q_{\text{loss}}(T_{\text{ambient}}), Q_{\text{reboiler}}) = \mathcal{M}(T_{\text{ambient}})$, where $\mathcal{M}(\cdot)$ is a model structure depending on certain parameters, Q_{reboiler} is a duty of the reboiler, which in this case is constant. As mentioned before, to find an accurate model of the distillation column, it is important to take the ambient temperature into account. This can be done by letting the estimated time-delay and the model parameters vary with ambient temperature, for example, by approximating that relationship between the model parameters, i.e. time-delays, zeros and poles, and the ambient temperature by low-order polynomials.

After performing various simulations, it became evident that it is not always possible to find a good relationship between the model parameters and the ambient temperature. The variations between the model parameters for different ambient temperatures were too large. For instance, ambient temperature variations between winter and summer periods can be quite large. Therefore, also the heat loss of the column can differ a great deal, and, as a consequence, the performance of the column.

We observe that the operation of the column becomes more or less constant when compensating for the heat losses due to ambient temperature, see Fig. 1. The heat losses are compensated by adjusting the reboiler power such that the difference between the power in the reboiler and the heat loss over the column is constant with respect to the variations in the ambient temperature, see Fig. 2. The difference between the reboiler power and the heat losses is also called the effective separation heat. It is defined by ΔQ, i.e. $Q_{\text{reboiler}}(T_{\text{ambient}}) - Q_{\text{loss}}(T_{\text{ambient}}) = \Delta Q$, throughout this paper. Thus, we use a reparametrization of the model, from $\mathcal{M}(Q_{\text{loss}}(T_{\text{ambient}}), Q_{\text{reboiler}})$ with Q_{reboiler} constant to $\mathcal{M}(Q_{\text{reboiler}}(T_{\text{ambient}}) - Q_{\text{loss}}(T_{\text{ambient}})) = \mathcal{M}(\Delta Q)$ with ΔQ constant, or independent of the ambient temperature T_{ambient}. Hence, it becomes possible to extract more accurate system models.

Thus, the proposed solution in this paper is to let the parameters, time-delays τ_i, zeros z_i and poles p_i, of the distillation

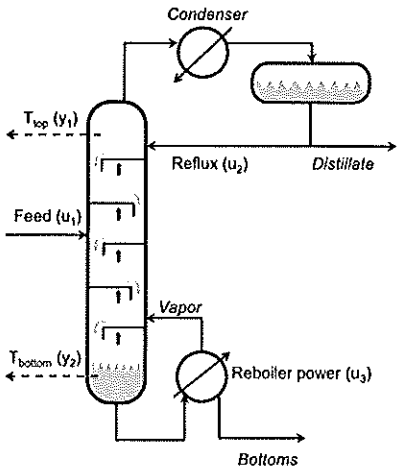


Fig. 3. A schematic representation of a binary distillation column. Here we consider three inputs: feed flow, u_1, reflux flow, u_2, and reboiler power, u_3, and two outputs: temperature at the top, y_1, and temperature at the bottom of the column, y_2.

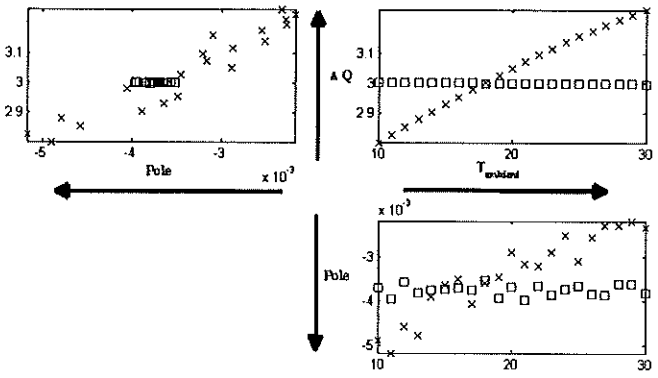


Fig. 4. Plot of one of the poles of the parametric model from reboiler power u_3 to the temperature at the bottom y_2: (×) ΔQ varies with the ambient temperature, (□) ΔQ is constant. Above left: relation between pole and ΔQ; above right: estimated ΔQ's depending on the ambient temperature (this plot is related to Fig. 2); bottom: relation between the pole and the ambient temperature.

column model depend on the heat inside the column ΔQ:

$$\hat{G}(s,\Delta Q) = \exp\left(-\tau\left(\Delta Q\right)s\right)\frac{\prod_{i=1}^{n_z}\left(s - z_i\left(\Delta Q\right)\right)}{\prod_{i=1}^{n_p}\left(s - p_i\left(\Delta Q\right)\right)},$$

where $\hat{G}(s,\Delta Q)$ denotes an estimate of the FRM in continuous Laplace domain s and where n_z and n_p denote the amount of zeros and poles, respectively.

III. Simulation results

This section shows the modeling results on a white-box model of a binary distillation column ([5], [6]) using frequency domain identification methods. The column, see Fig. 3, consists of 38 stages: a reboiler, 36 trays and a condenser. In this project, we consider a MIMO distillation system with: three inputs (the reboiler power, feed flow into the column, and reflux flow) and two outputs (temperatures in the reboiler and at the top of the column).

The model of the distillation column is extracted for a constant ambient temperature level from 10 to 30 °C, $T_{\text{ambient}} \in [10, 30]$ °C. For each ambient temperature, two realizations of the system's response are simulated and within each realization there are three experiments with orthogonal multisine excitations. The transient responses of these simulations are processed with the robust LPM to get a nonparametric estimate of the FRM together with its (total) variance bounds. In the final step, the nonparametric FRM and its variance bounds are used to find a parametric rational form Laplace model of maximum order four and with a time-delay term.

The frequency domain continuous-time Laplace models are compared for cases when ΔQ varies and when ΔQ is constant. Fig. 4 shows one of the poles of the third order rational model from reboiler power u_3 to the temperature at the bottom y_2. The fitted parametric model has zero delay. Fig. 4 shows the relations between the pole and ΔQ, between ΔQ's and the ambient temperature, and between the pole and the ambient temperature. The plot in the above right corner is related to Fig. 2, except that the actual difference between the reboiler power and the heat loss over the column, ΔQ, is shown. For a constant reboiler power, ΔQ varies with the ambient temperature and consequently, the pole varies with the ambient temperature. This is shown in Fig. 4 at the bottom (marked with a cross). When the reboiler power is adjusted such that the difference between the power in the reboiler and the heat loss over the column is independent of the ambient temperature, i.e. ΔQ is contant, the pole as a function of ambient temperature becomes more or less constant, see Fig. 4 at the bottom (marked with a square). This means that the model of the distillation column becomes almost independent of the ambient temperature. We also observe this in the top left plot of Fig. 4: for ΔQ varying with ambient temperature, the pole also varies; for constant ΔQ, the pole stays between $-4\cdot10^{-3}$ and $-3.5\cdot10^{-3}$ Hz for all the ambient temperatures. Consequently, it becomes easier to estimate a meta-model of the distillation column whose parameters depend on the effective separation heat ΔQ.

IV. Measurement results

The experimental setup involves a computer-controlled packed distillation column. The column is about 6 m high and has an internal diameter of 6 cm. The column works under atmospheric conditions and contains three sections of about 1.5 m with Sulzer CY packing (Sulzer, Winterthur) responsible for the separation. This packing has a contact surface of 700 m^2 per 1 m^3 of column's volume and each meter packing is equivalent to 3 theoretical trays. The feed stream containing a mixture of methanol and isopropanol is

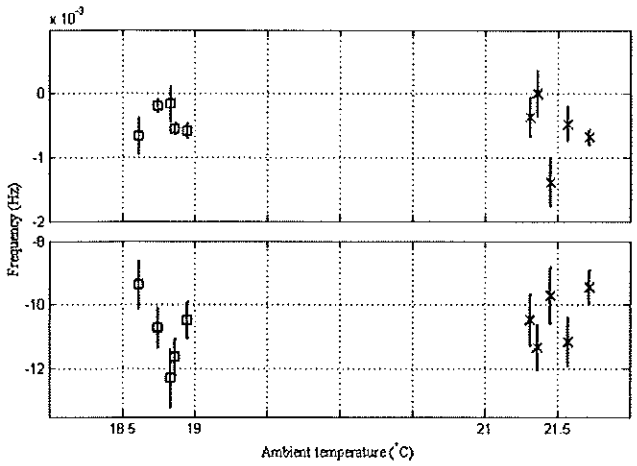


Fig. 5. Plot presents poles of a second order parametric model from reboiler power to temperature at the bottom using real measurements. The data is divided into intervals of two periods with 50% overlap. The poles of the extracted models, together with their 68%-confidence bounds, are plotted against the mean of the ambient temperature during an interval. (×) Variations of heat loss due to the ambient temperature are not compensated; (□) heat loss is compensated in the reboiler power for the ambient temperature variations and for the difference between heat losses of both experiments. (red) Low-frequency poles; (black) high-frequency poles.

fed into the column between packed sections 2 and 3. The temperature of the feed can be adjusted by an electric heater of maximum 250 W. At the bottom of the column a reboiler is present containing two electric heaters of maximum 3000 W each. In the reboiler, a part of the liquid is vaporised while the rest is extracted as bottom stream. At the top of the column a total condenser allows the condensation of the entire overhead vapour stream, which is then collected in a reflux drum. A part of the condensed liquid is fed back to the column as reflux, while the remainder leaves the column as the distillate stream. All temperatures are measured in degrees Celsius. The actuators and sensors are connected to a Compact Fieldpoint (National Instruments, Austin) with a controller interface cFP-2100 and I/O modules cFP-AIO-610, cFP-AIO-610 and cFP-AI-110. A Labview (National Instruments, Austin) program has been developed to control the actuators and to register the variables.

In this setup we manipulated only the reboiler power, u_3 in Fig 3. The input on the reboiler is a random multisine excitation, with a period of one hour. The other two variables, feed and reflux, were kept at a constant level. Measurements are available for the temperature at the bottom, y_2, and the temperature at the top, y_1. The ambient temperature is measured one meter from the column. The sampling period of all the data is 0.1 second.

Two experiments are done. Experiment 1: the heat losses and thus also ΔQ are varying with ambient temperature. Experiment 2: the difference in the heat losses with respect to Experiment 1 and also the variations of the heat losses with respect to the variations of ambient temperature are compensated in first order in the reboiler power. In this way, ΔQ of Experiment 2 is, on average, at the same level as ΔQ from Experiment 1. Also, ΔQ of Experiment 2 is kept constant with respect to the variations in ambient temperature.

Contrary to the simulation, we can not manipulate the ambient temperature. To overcome this problem, the measurement data of both experiments is divided and processed per two periods (by two hours). For each two periods we estimate a nonparametric FRF model together with its error bounds using the robust LPM with 1 experiment and 1 realization (see Subsection II-B). Then, using these estimates we find a low-order parametric FRF. From both measurements of 6 hours long, we got 5 intervals of data per experiment by using 50% overlap between two successive intervals.

As in Section III, we estimate an FRF from the reboiler power u_3 to the temperature at the bottom y_2. A second order FRF without delay is sufficient for both experiments. Fig. 5 shows both low- and high-frequency poles with their 68%-confidence bounds of the fitted FRF's. The ambient temperature difference between the two experiments is, on average, more than two degrees. During the individual experiments the ambient temperature variations were less than one degree, as we can see from Fig. 5. Squares indicate the poles from Experiment 2 (where the heat losses were compensated) for all the two-periods data intervals. Crosses indicate the poles from Experiment 1.

Fig. 5 shows that the high-frequency poles (black) have relatively large confidence bounds for both experiments. The low-frequency poles (red) have smaller confidence bounds. This plot shows that low-frequency pole for the compensated heat loss (Experiment 2) varies less with the ambient temperature than the one for the noncompensated heat loss (Experiment 1). Comparing the poles' magnitudes between Experiment 1 and Experiment 2 shows that the poles of Experiment 2 are more or less of the same magnitude as poles of Experiment 1. The measurement results presented here are not as conclusive as the simulation results in Section III. For comparison, see bottom plot of Fig. 4. An explanation for these differences could be, that the ambient temperature did not change much during the measurement experiments. Thus, more experiments with larger ambient temperature variations are needed to confirm the results from Section III.

V. Conclusion and Discussion

In this paper we have applied a recently developed robust local polynomial method. Using this method we can obtain a good nonparametric estimate of the FRM of our system from the transient response of the system to a random-phase multisine excitation. In addition, the local polynomial method makes it possible to estimate the variance of the nonparametric FRM in the presence of both nonlinearities and additive noise. Furthermore, we have identified a parametric model with parameters depending on the difference between the reboiler power and the heat loss of the column, ΔQ. We found that the operating point of the column is more or less the same when keeping ΔQ constant with respect to the variations in the

ambient temperature. The results were illustrated using both simulation data and using real measurements of a pilot-scale binary distillation column at KULeuven-BioTeC. Thus, by keeping the operating point of the system constant, it becomes easier to estimate the relation between the system and the ambient temperature. Future work will contain an extraction of a meta-model of the distillation column whose parameters depend on the effective separation heat varying with ambient temperature.

ACKNOWLEDGMENT

This work is sponsored by the Vrije Universiteit Brussel, dept. ELEC, Pleinlaan 2, 1050 Brussels, Belgium; Fund for Scientific Research (FWO-Vlaanderen) project FWOAL561; the Belgian Federal Government (IUAP VI/4); OT/10/035; Optimization in Engineering (OPTEC) PFV/10/002; SCORES4CHEM KP/09/005; and Belgian Program on Interuniversity Poles of Attraction, initiated by the Belgian Federal Science Policy Office.

REFERENCES

[1] S. Skogestad, "Dynamics and control of distillation columns - a critical survey," *Modeling, identification and control*, vol. 18, no. 3, pp. 177–217, 1997.

[2] D. E. Rivera, H. Lee, H. D. Mittelman, and M. W. Braun, "High-purity distillation: using plant-friendly multisine signals to identify a strongly interactive process," *IEEE Control Systems*, vol. 27, no. 5, pp. 72–89, 2007.

[3] E. W. Jacobsen and S. Skogestad, "Identification of ill-conditioned plants - a benchmark problem," in *The Modeling of Uncertainty in Control*. Springer, 1994, pp. 367–376.

[4] R. Pintelon and J. Schoukens, *System identification. A frequency domain approach*. IEEE, 2001.

[5] M. Diehl, "Real-time optimization for large scale nonlinear processes," Ph.D. dissertation, University of Heidelberg, 2002.

[6] J. Bonilla, "Structure and convexity exploitation in nonlinear chemical process modeling and estimation," Ph.D. dissertation, Katholieke Universiteit Leuven, 2011.

[7] T. P. Dobrowiechki, J. Schoukens, and P. Guillaume, "Optimized excitation signals for MIMO frequency response function measurements," *IEEE Transactions on Instrumentation and Measurement*, vol. 55, no. 6, pp. 2072–2079, 2006.

[8] R. Pintelon, J. Schoukens, G. Vandersteen, and K. Barbé, "Estimation of nonparametric noise and FRF models for multivariable systems – part I: theory," *Mechanical Systems and Signal Processing*, vol. 24, no. 3, pp. 573–595, 2010.

[9] R. Pintelon, G. Vandersteen, J. Schoukens, and Y. Rolain, "Improved (non-)parametric identification of dynamic systems excited by periodic signals - the multivariate case," *Mechanical Systems and Signal Processing*, vol. 25, no. 8, pp. 2892–2922, 2011.

[10] "Frequency domain system identification toolbox for MATLAB - general information." [Online]. Available: http://matlab.gamax.hu/english/products/fdident/