Nonlinear identification and control of Organic Rankine Cycle systems using sparse polynomial models.

Andres Hernandez[1,3], Fredy Ruiz[2], Adriano Desideri[3], Clara Ionescu[1], Sylvain Quoilin[3],
Vincent Lemort[3] and Robin De Keyser[1]

Abstract— Development of a first principles model of a system is not only a time- and cost- consuming task, but often leads to model structures which are not directly usable to design a controller using current available methodologies. In this paper we use a sparse identification procedure to obtain a nonlinear polynomial model. Since this is a NP-hard problem, a relaxed algorithm is employed to accelerate its convergence speed. The obtained model is further used inside the nonlinear Extended Prediction Self-Adaptive control (NEPSAC) approach to Nonlinear Model Predictive Control (NMPC), which replaces the complex nonlinear optimization problem by a simpler iterative quadratic programming procedure. An organic Rankine cycle system, characterized for presenting nonlinear time-varying dynamics, is used as benchmark to illustrate the effectiveness of the proposed combined strategies.

I. INTRODUCTION

Today process industry is experiencing an increasing productivity demand. This, together with tighter environmental regulations, demanding economical considerations and the need of operating in an energy-efficient manner, is forcing industry to operate systems closer to the boundary of the admissible operating region, where productivity is often maximized. In this regard modeling and control design play an important role to achieve optimal operation of a system.

One of the control strategies which has been well-accepted by industry is Model Predictive Control (MPC) [1]. MPC refers to a family of control approaches, which makes explicit use of a model of the process to optimally obtain the control signal by minimizing an objective function [2]. Due to this active use of a model, its performance is strongly linked to the model's ability to accurately predict the system dynamics. Real processes are inherently nonlinear, thus ideally a low-order nonlinear model and therefore a nonlinear MPC should be used to control the system. However, linear predictive controllers have been predominantly used, this is due to the easy to obtain linear parametric models compared to nonlinear ones [3].

Many model classes have been considered for nonlinear identification purposes, e.g., Wiener or Hammerstein, neuronal networks, radial basis functions, support vector machines and polynomial expansions. From those the polyno-mial nonlinear autoregressive exogenous (NARX) models are particular interesting, since they allow a direct interpretation and analysis of the nonlinear input-output dependencies. Several experimental studies have shown that many systems can be accurately approximated by polynomial NARX models of relatively small size (see e.g., [4], [5]).

In this contribution the objective is to identify a sparse nonlinear polynomial model using a modified version of the Elastic Net regression algorithm originally proposed in [6]. First it is essential to clarify what sparse means in this context. It consists in approximating a function using a 'few' basis functions properly selected within a 'large' set. More precisely, a sparse approximation is a linear combination of fixed basis functions, where the vector of linear combination coefficients is sparse, i.e. it has only a 'few' non-zero elements. The pioneer work performed in [6], [7] shows that a sparse model can be obtained by making a basic augmentation of the ordinary least squares solution, with the goal of obtaining a coefficient shrinkage and selection. The augmentation often consists in including penalty functions to the least square problem, thus providing some regularization. In [6], [7] and the references therein, it is shown that for sufficient amounts of regularization, solutions are sparse, i.e., some of the coefficients of the model are exactly zero, leading to more compact models which are easier to interpret.

The modification here proposed in section II, is used to accelerate convergence speed by relaxing the optimization problem, whereas the goal is to minimize a combination of the ℓ_1- and ℓ_2- norm of the parameter vector, with a constraint on the model accuracy. The relaxed Sparse identification algorithm allows to obtain the structure of the model, then a second step is implemented by identifying its coefficients using a least-squares criterion and applying pruning. Finally an input/output polynomial model of the system with prescribed modeling error is obtained [8]. The methodology is similar to the one known as simulation error minimization with pruning (SEMP), which as reported in [9], can yield an increase in robustness for model selection compared to prediction error minimization (PEM) framework with respect to low excitation identification conditions.

Then the obtained polynomial model is used to construct a nonlinear MPC strategy, i.e., the Nonlinear Extended Prediction Self-Adaptive Control (NEPSAC) approach to NMPC. This controller, fully described in [10], can deal with models in different formats (e.g. state-space, transfer functions, neuronal networks, polynomial parametric models, etc.) as it uses the nonlinear model for prediction directly,

[1]A. Hernandez, C. Ionescu and R. De Keyser are with Faculty of Engineering and Architecture, Department of Electrical energy, Systems and Automation, Ghent University, Belgium. `andres.hernandez@ugent.be`
[2]F. Ruiz is with Department of electronics engineering, Pontificia Universidad Javeriana, Colombia. `ruizf@javeriana.edu.co`
[3]S. Quoilin and V. Lemort are with Thermodynamics Laboratory, University of Liege, Belgium. `squoilin@ulg.ac.be`

taking full advantage of the given nonlinear system dynamics to generate a high-performance design without involving model linearization or gain scheduling for its implementation. Additionally, the usually complex nonlinear optimization problem, is in the NEPSAC algorithm replaced by a simpler iterative quadratic programming procedure, to solve problems including constraints as briefly described in section III.

The effectiveness of the proposed combined strategy is tested in simulation using an organic Rankine cycle (ORC) system for waste heat recovery applications (WHR). This system is particularly challenging, since its dynamics are highly nonlinear and time-varying because they change as function of the waste-heat source [11]. Their dynamic components are mainly allocated in the evaporator, therefore research in modeling has been focused on finding a trade-off between model accuracy and simulation speed as evaluated in [12], where finite volume and moving boundary approaches to evaporator modeling are proposed.

Regarding the control design for ORC systems, it is essential to guarantee a vapor quality at the inlet of the expander, since liquid droplets can damage the machine. This requirement can be achieved by ensuring a certain level of superheating [13], [14]. For example, in [15] a feed-forward and gain-scheduled PI controller is proposed. In [16] an explicit multi-model MPC is applied to control superheating. In conclusion, identifying a nonlinear model for superheating and achieving a proper regulation of this signal despite waste-heat variations represents the target for the controller.

The performance of the nonlinear identification and control algorithms is evaluated by developing a superheating control of an ORC system as described in section IV. Finally, the main outcome of this investigation is summarized in section V.

II. SPARSE IDENTIFICATION OF NONLINEAR FUNCTIONS

In cases where an accurate first principle model is not available for a complex system, models deduced from experimental data by means of system identification techniques provide reliable alternatives. Whilst existing system identification methods for linear systems can be considered to be at a mature stage, nonlinear system identification techniques still reveal several fundamental issues to be resolved to further improve general model performance. In this section a sparse identification algorithm of nonlinear functions is presented.

Identifying a sparse approximation of a function from a set of data, possibly corrupted by noise, is what is called sparse identification. It consists in constructing a sparse model which involves a minimum model variables and terms but can generate a maximum of model generalization performance. Therefore, a sparse approximation is a linear combination of fixed basis functions, where the vector of linear combination coefficients is sparse, i.e. it has only a 'few' non-zero elements.

A. Problem formulation

Consider a nonlinear function f_0 defined by

$$y = f_0(x) \tag{1}$$

where $x \in \mathbb{R}^{n_x}$, $y \in \mathbb{R}$. Suppose that f_0 is not known but a set of noise-corrupted data $D = \{\tilde{x}, \tilde{y}\}_{k=1}^{L}$ is available, described by

$$\tilde{y}_k = f_0(\tilde{x}_k) + n_k, \quad k = 1, 2, \ldots, L \tag{2}$$

where n_k is noise. Define the following parametrized function:

$$f_a(x) = \sum_{i=1}^{n} a_i \phi_i(x) = \phi(x)a \tag{3}$$

where $\phi(x) = [\phi_1(x), \phi_2(x), \ldots, \phi_n(x)]$, $\phi_i : X \to Y$ are known fixed basis functions, and $a = [a_1, a_2, \ldots, a_n]^T \in \mathbb{R}^n$ is an unknown coefficient vector.

Problem 1: from the data set D, identify a coefficient vector a such that

(i) is sparse and
(ii) the identification error

$$e(f_a) \doteq \|f_0 - f_a\|_2 \tag{4}$$

is small and bounded.

Sparse solutions may be preferred to full counterparts if the model can be assumed to be sparse or when interpretation of the results can be important.

B. Sparse Identification algorithm

The sparsity of a vector is typically measured by the ℓ_0 quasi-norm, defined as the number of its non-zero elements. Sparse identification can be performed by looking for a coefficient vector of the basis function linear combination with a 'small' ℓ_0 quasi-norm, that yields a 'small' prediction error ε between the measured output $\tilde{y}$ and the predicted output $f_a(\tilde{x}) = \phi(\tilde{x})a$. Thus, a solution to problem 1 could be found by solving the following optimization problem:

$$a^0 = arg \min_{a \in \mathbb{R}^n} \|a\|_0$$
$$subject\ to\ \|\tilde{y} - \Phi a\|_2 \leq \varepsilon \tag{5}$$

with ε the maximum desired prediction error and

$$\tilde{y} \doteq (\tilde{y}_1, \ldots, \tilde{y}_L)$$
$$\Phi \doteq \begin{bmatrix} \phi_1(\tilde{x}_1) & \cdots & \phi_n(\tilde{x}_1) \\ \vdots & \ddots & \vdots \\ \phi_1(\tilde{x}_L) & \cdots & \phi_n(\tilde{x}_L) \end{bmatrix} \tag{6}$$
$$\doteq [\phi_1(\tilde{x}) \ldots \phi_n(\tilde{x})]$$

However, the ℓ_0 is a non-convex function and therefore minimization of (5) is in general an NP-hard problem. Similarly as proposed by [8], we propose a convex relaxation, where an optimization problem similar to (5) is solved, by replacing the ℓ_0 quasi-norm by a combination of a ℓ_1 and weighted ℓ_2 norms. The optimization problem (5) is thus redefined as a two-step algorithm:

Step 1: solve the optimization problem

$$a^1 = arg \min_{a \in \mathbb{R}^n} \quad L^{-1} \|a\|_1 + \beta \|a\|_2$$
$$s.t. \|\tilde{y} - \Phi a\|_2 \leq \varepsilon \tag{7}$$

with L the length of the data set, β a weighting coefficient and ε the maximum expected prediction error.

This structure of the optimization problem in equation (7) is chosen, as the ℓ_2 penalty ensures an unique solution also when the number of basis functions is larger than the data length, and the ℓ_1 penalty offers variable selection via a sparse vector of coefficients a as demonstrated in [6].

Step 2: define a sub-space with the active basis of a^1 and compute the estimate a^* of a^0 using least-squares of prediction error

$$a^* = arg \min_{a_i \in \mathbb{R}^n} \|\tilde{y} - \Phi a_i\|_2$$
$$s.t. a_i = 0, \forall i \, |a_i^1| < \gamma \tag{8}$$

with γ an additional tolerance value used for pruning, thus enhancing sparsity.

Note that the major problem with regularization methods is the optimal selection of the regularization parameters, in this case γ, β and ε, which correspond to a complexity/bias vs. variance trade-off [6].

III. NMPC-NEPSAC ALGORITHM

A brief introduction to NEPSAC algorithm is presented in this section. For a detailed description the reader is referred to [10] and [17].

A. Computing the Predictions

Using NEPSAC algorithm, the measured process output can be represented as:

$$y(t) = x(t) + n(t) \tag{9}$$

where $x(t)$ is the model output which represents the effect of the control input $u(t)$ and $n(t)$ represents the effect of the disturbances and modeling errors, all at discrete-time index t. Model output $x(t)$ can be described by the generic system dynamic model:

$$x(t) = f[x(t-1), x(t-2), \ldots, u(t-1), u(t-2), \ldots] \tag{10}$$

Notice that $x(t)$ represents here the model output, not the state vector. Also important is the fact that f can be either a *linear* or a *nonlinear* function.

Furthermore, the disturbance $n(t)$ can be modeled as colored noise through a filter with the transfer function:

$$n(t) = \frac{C(q^{-1})}{D(q^{-1})} e(t) \tag{11}$$

with $e(t)$ uncorrelated (white) noise with zero-mean and C, D monic polynomials in the backward shift operator q^{-1}. The disturbance model must be designed to achieve robustness of the control loop against unmeasured disturbances and modeling errors [2].

A fundamental step in the MPC methodology consists of the prediction. Using the generic process model (9), the predicted values of the output are:

$$y(t+k|t) = x(t+k|t) + n(t+k|t) \tag{12}$$

$x(t+k|t)$ and $n(t+k|t)$ can be predicted by recursion of the process model (10) and by using filtering techniques on the noise model (11), respectively [10].

B. Optimization Procedure

A key element in linear MPC is the use of base (or free) and optimizing (or forced) response concepts [2]. This is however no longer valid for nonlinear processes, since the superposition principle does not hold in that case. In NEPSAC this is solved as follows:

The future response can then be expressed as:

$$y(t+k|t) = y_{base}(t+k|t) + y_{optimize}(t+k|t) \tag{13}$$

The two contributing factors have the following origin:

- $y_{base}(t+k|t)$ is the effect of the past inputs, the apriori defined future base control sequence $u_{base}(t+k|t)$ and the predicted disturbance $n(t+k|t)$.
- $y_{optimize}(t+k|t)$ is the effect of the additions $\delta u(t+k|t)$ that are optimized and added to $u_{base}(t+k|t)$, according to $\delta u(t+k|t) = u(t+k|t) - u_{base}(t+k|t)$. The effect of these additions is the discrete time convolution of $\Delta U = \{\delta u(t|t), \ldots, \delta u(t+N_u-1|t)\}$ with the impulse response coefficients of the system (G matrix), where N_u is the chosen control horizon.

The control ΔU is the solution to the following constrained optimization problem:

$$\Delta U = arg \min_{\Delta U \in \mathbb{R}^{N_u}} \sum_{k=N_1}^{N_2} [r(t+k|t) - y(t+k|t)]^2$$
$$s.t. A.\Delta U \leq B \tag{14}$$

where N_1 and N_2 are the minimum and maximum prediction horizons, N_u is the control horizon, $r(t+k|t)$ is a future setpoint or reference sequence. The various process inputs and output constraints can all be expressed in terms of ΔU, resulting in matrices A, B.

As the cost function (14) with decision variables ΔU is quadratic with linear constraints, then the minimization problem can be solved by a QP algorithm.

C. NEPSAC iterative procedure

The procedure described until now can be used to implement a linear MPC (EPSAC), if $f[\cdot]$ in (10) is linear, as superposition principle holds in (13). Instead, if a nonlinear model is employed, the condition is that the term $y_{optimize}(t+k|t)$ is small enough compared to $y_{base}(t+k|t)$, in order to make the superposition principle still valid. Such condition can be satisfied if $\delta u(t+k|t)$ is close to zero, making thus $u_{base}(t+k|t)$ close to the optimal $u^*(t+k|t)$. To address this issue, the idea is to recursively compute $\delta u(t+k|t)$ solving (14), within the same sample instant, until $\delta u(t+k|t)$ converges to 0. Inside the recursion $u_{base}(t+k|t)$ is updated each time to $u_{base}(t+k|t) + \delta u(t+k|t)$.

Also important is the fact that in NEPSAC the G matrix has to be recomputed at every sampling instant [17]. The number of required iterations depends on how far the optimal $u^*(t+k|t)$ is away with respect to $u^*(t+k|t-1)$. In

quasi-steady-state situations, the number of iterations is low (1 . . . 2) [17]. On the other hand, during transients the number of iterations might raise to 10 [17].

IV. CASE STUDY: WASTE HEAT RECOVERY APPLICATION

This section describes the architecture and main characteristics of the Organic Rankine Cycle (ORC) system used for evaluating the performance of the developed control strategies.

A. The Organic Rankine Cycle system

The ORC power unit investigated in this work is based on a sub-critical $11\,kW_{el}$ unit for stationary low temperature waste heat recovery. The system is based on a regenerative cycle and Solkatherm (SES36) as working fluid. The $11\,kW_{el}$ expander is originally a single screw compressor adapted to run in expander mode.
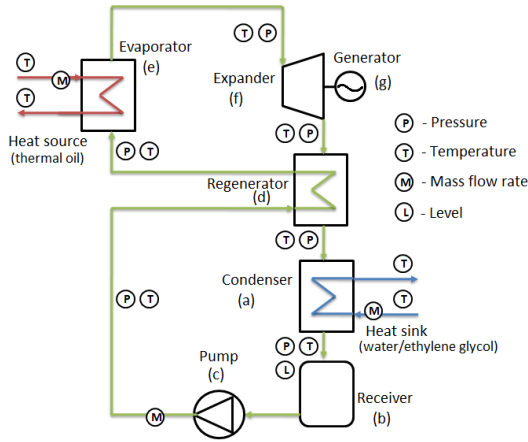

Fig. 1. Schematic layout of the organic Rankine cycle power unit.

A schematic layout of the ORC system is presented in Fig. 1. Starting from the bottom of the scheme, it is possible to recognize the liquid receiver (b) installed at the outlet of the condenser (a) where the fluid is collected in saturated liquid condition. From the receiver outlet, the fluid is pumped (c) through the regenerator (d) cold side, and the evaporator (e), where it is heated up to superheated vapor, reaching its maximum temperature at the evaporator outlet. The fluid, after being expanded in the volumetric machine (f), enters the regenerator hot side, and then it flows into the condenser (a) to close the cycle.

In order to assess the performance of the different developed control strategies, a validated dynamic model of the ORC system presented in Fig. 1 has been developed in the Modelica language [12], using existent components from the ThermoCycle library [18]. The developed Modelica model is then exported into Simulink/Matlab environment by means of the Functional Mock-Up Interface (FMI) open standard, using a model exchange format. This simulation approach takes advantage of the strengths of each platform: Modelica for modeling and Simulink/Matlab for control design. For the simulations performed in this paper, the generator rotational

speed is kept constant at 3000 rpm to emulate an installation directly connected to the grid.

Regarding the optimal operation of the ORC unit, the superheating and the evaporating temperature are the most relevant variables to be controlled [19]. The superheating is defined as:

$$\Delta T_{sh} = T_{exp,su} - T_{sat,ev} \tag{15}$$

where $T_{exp,su}$ is the temperature measured at the inlet of the expander and $T_{sat,ev}$ the evaporating temperature, corresponding to the temperature at which the fluid undergoes the phase transition from saturated liquid to saturated vapor at the fixed evaporating pressure $P_{sat,ev}$.

Research performed on ORC technology (e.g., [13], [14]) has already established that to optimally operate the ORC power unit, two conditions need to be satisfied: **1)** A 'high' efficiency and a safe operation of the ORC unit is achieved if a minimum amount of superheating at expander inlet is guaranteed. **2)** For each heat source condition there exists an optimal evaporating temperature which maximizes the output power. In this study we will focus on fulfilling the first condition, by designing a controller for superheating. In order to assess the overall performance of the system the net-output power $\dot{W}_{el,net}$ is computed:

$$\dot{W}_{el,net} = \dot{W}_{exp} - \dot{W}_{pump} \tag{16}$$

where $\dot{W}_{exp}$ is the expander electrical power, $\dot{W}_{pump}$ is the pump electrical power.

B. Low-order model suitable for prediction

The ORC unit considered has one manipulated variable (the pump speed N_{pp}), and two outputs (the evaporating temperature $T_{sat,ev}$ and the superheating ΔT_{sh}). Notice that the temperature T_{hf} and mass flow variations m_{hf} in the heat source also influence $T_{sat,ev}$ and ΔT_{sh}, thus becoming two measured disturbances. As result we are interested on identifying a system consisting of 3 inputs (one manipulated and two measured disturbances) and 2 outputs.

The following input signals have been considered to perform the identification: the heat source conditions T_{hf} and m_{hf} have been simulated as white noise filtered to a maximum band of 0.5 rad/s and 0.63 rad/s with amplitude variations of $\pm 20°C$ around $110°C$ and $\pm 0.5\,kg/s$ around $1.0\,kg/s$, respectively. The pump speed has been taken as the sum of 10 sinusoids spread over the band $[0.005, 0.63]\,rad/s$, taking values between 1320 rpm and 2100 rpm. A set of $L = 2000$ samples has been generated from the 'true' system, considering a sampling time of $T_s = 1\,s$.

A *linear parametric identification* is performed, using the prediction error minimization method in the data collected from multisine excitation signals [20]. The identified linear model in the form of discrete-time transfer functions for a sampling time of $T_s = 1\,s$, is presented in (17):

$$\frac{\Delta T_{sh}(q)}{N_{pp}(q)} = \frac{-0.063q^{-1}+0.059q^{-2}}{1-2.44q^{-1}+1.955q^{-2}-0.51q^{-3}} \tag{17a}$$

$$\frac{\Delta T_{sh}(q)}{T_{hf}(q)} = \frac{0.47q^{-1}}{1-0.51q^{-1}} \tag{17b}$$

$$\frac{\Delta T_{sh}(q)}{m_{hf}(q)} = \frac{-2.98q^{-1}+4.29q^{-2}-1.31q^{-3}}{1-1.35q^{-1}-0.11q^{-2}+0.46q^{-3}} \tag{17c}$$

$$\frac{T_{sat,ev}(q)}{N_{pp}(q)} = \frac{0.066q^{-1}-0.063q^{-2}}{1-2.42q^{-1}+1.91q^{-2}-0.49q^{-3}} \tag{17d}$$

$$\frac{T_{sat,ev}(q)}{T_{hf}(q)} = \frac{0.0017q^{-11}-0.0017q^{-12}}{1-3.6q^{-1}+4.88q^{-2}-2.95q^{-3}+0.67q^{-4}} \tag{17e}$$

$$\frac{T_{sat,ev}(q)}{m_{hf}(q)} = \frac{2.43q^{-1}-6.16q^{-2}+5.33q^{-3}-1.6q^{-4}}{1-2.93q^{-1}+3.12q^{-2}-1.42q^{-3}+0.23q^{-4}} \tag{17f}$$

Fig. 2. Prediction of superheating using the identified linear and non-linear models as simulators.

Next, we are interested in identifying a *nonlinear polynomial model* of the system. In order to perform the sparse identification, a set of $n = 57$ polynomial basis functions has been considered and the corresponding matrix $\Phi = (\Phi_1(\tilde{x}),\ldots,\Phi_L(\tilde{x}))$ has been obtained according to:

$$
\begin{aligned}
\Phi_k(\tilde{x}) =& [\phi_1(\tilde{x}),\ldots,\phi_n(\tilde{x}))] \\
=& [1, y(k-r), N_{pp}(k-r), T_{hf}(k-r), m_{hf}(k-r), \\
& y(k-r)^2, y(k-r)*N_{pp}(k-r), y(k-r)*T_{hf}(k-r), \\
& y(k-r)*m_{hf}(k-r), N_{pp}(k-r)^2, \\
& N_{pp}(k-r)*T_{hf}(k-r), N_{pp}(k-r)*m_{hf}(k-r), \\
& T_{hf}(k-r)^2, T_{hf}(k-r)*m_{hf}(k-r), m_{hf}(k-r)^2]
\end{aligned} \tag{18}
$$

where r is the regressor which defines how many variables we look in the past, for example $r = 2$ represents $y(k-r) \equiv [y(k-1), y(k-2)]$; in this example $r = 4$ and $k = 1, 2, \ldots, L$.

Using the regularization parameters $\beta = 20$, $\gamma = 1e-6$ and $\varepsilon = 1\,^\circ\text{C}$, the identified model for ΔT_{sh} requires of 17 active basis functions. Simulations using the validation data are performed for both linear and nonlinear models (Fig. 2). An accurate identification of low superheating values is still a challenging task, modeling errors can be attributed to the choice of the fixed basis functions, these results suggest that other type of nonlinearities should be considered. The normalized root-mean-square error (RMSE,[20]) of each model to the validation data is computed, reporting that a fitting percentage of 48% and 72% for the linear and nonlinear model, respectively. Nevertheless, from control point of view, this is not a problem as still a robust controller can be tuned to deal with modeling errors or unmodeled dynamics.

C. Application to constrained predictive control

Two predictive controllers are implemented, a EPSAC-MPC based on the linear parametric model (17) and a NEPSAC-NMPC based on the nonlinear polynomial model. The controllers are tested in simulation using a heat source profile which could be typically observed in industrial waste heat [13], as for example from exhaust gases from a reheat furnace, as depicted in Fig. 3.

The performance of the controllers is assessed during a tracking test, where it is expected to have the fastest response with minimum overshoot, while satisfying actuator constraints ($N_{pp,min} = 1320\,\text{rpm}$; $N_{pp,max} = 2100\,\text{rpm}$; $\Delta N_{pp} =$
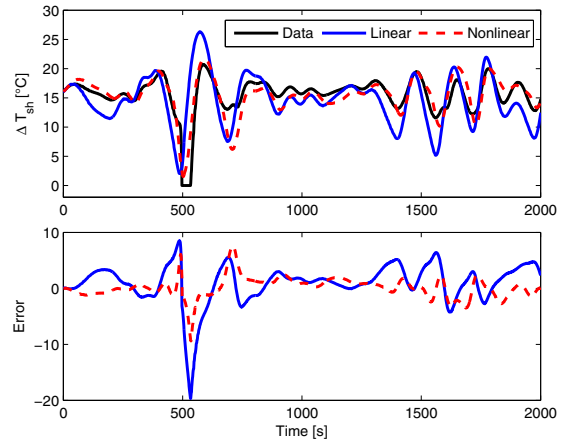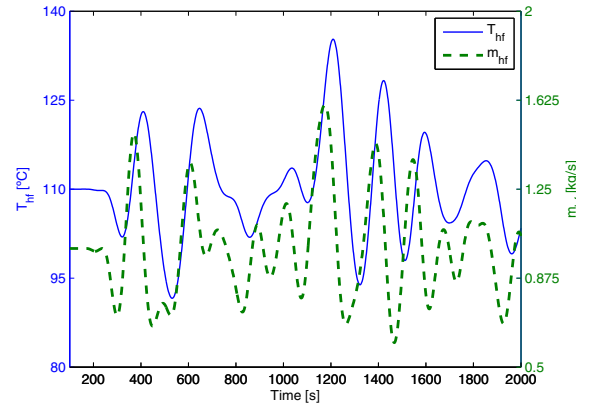

Fig. 3. Temperature and mass flow rate of the defined heat source.

$100\,\text{rpm}$) and constraints at the process output ($\Delta T_{sh,min} = 10\,^\circ\text{C}$). Both linear and nonlinear controllers are tuned for horizons $N_1 = 1$ and $N_2 = 10$.

The performance of the control strategies is depicted in Fig. 4, both linear and nonlinear controllers are able to track the reference, however the linear controller presents oscillations at low superheating values due to large modeling errors. The Integral Absolute Error (IAE) is used as criteria to quantify the control performance for tracking the reference, reporting as result 972.76 and 750.84, for the linear and nonlinear control, respectively. The control effort is similar in magnitude for both controllers, except for the observed oscillations, thus suggesting that either the region/conditions where the controller has been tested presents a weak nonlinearity or that the nonlinear model is still not capturing the main nonlinearity present in the system. From a practical point of view, the nonlinear controller achieves a smoother control effort which might be beneficial for the actuator (i.e., enlarging the pump life).
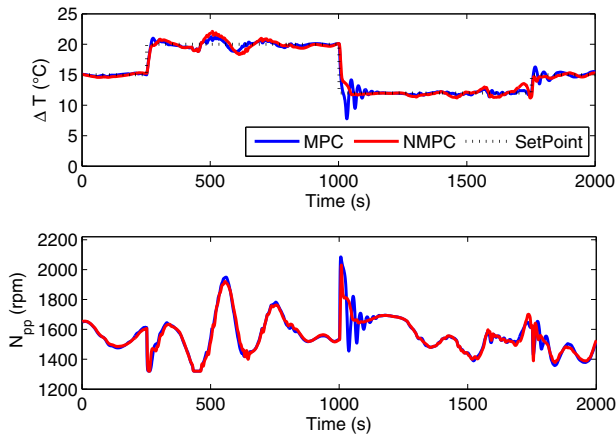
Fig. 4. Control performance of MPC and NMPC, for waste heat recovery application using an Organic Rankine Cycle.

V. CONCLUSIONS

In the present contribution, a relaxed sparse identification algorithm is presented in order to identify a parametric model of a nonlinear system, with prescribed modeling error. The augmentation of the basic ordinary least square solution with a ℓ_1 penalty allows to identify the structure of the model and to enhance sparsity, while the ℓ_2 penalty ensures obtaining an unique solution. Sparse polynomial models compared to other black-box identification techniques, for example Neuronal Networks, offer the benefit that a certain degree of understanding about the type (e.g., square, cubic, etc) and location (e.g., input, output, etc) of nonlinearity present in the system can be obtained. The simulation example provided in this paper is of industrial relevance and represents a challenging for modeling and control design. The optimal performance for these systems is often achieved, by proper regulation of the quality of vapor entering the expander. The latter can be observed by computing the level of superheating, which presents highly nonlinear and time-varying dynamics. Results show that it is possible to identify a sparse polynomial model that describes the dynamics over a wide operating range.

The obtained input/output polynomial model is further used in order to construct a constrained Nonlinear Model Predictive Control strategy using the NEPSAC approach. The performance of the proposed controller is tested and compared to the one achieved using a linear model predictive controller for the studied case. The obtained results suggest that the NMPC strategy leads to a smoother, safer and more efficient operation; where similar or better tracking performance is obtained at a lower control effort.

Future work includes identification and control of multiple input multiple output systems and using the knowledge of the prescribed modeling error to build a robust NMPC controller.

ACKNOWLEDGMENT

REFERENCES

[1] S. Qin and T. Badgwell., "A survey of industrial model predictive control technology." *Control engineering practice*, vol. 11, pp. 733–764, 2003.

[2] E. F. Camacho and C. Bordons., *Model Predictive Control*, 2nd ed. Springer-Verlag, London,, 2004, vol. 405 pages.

[3] S. A. Billings, *Nonlinear System Identification: NARMAX Methods in the Time, Frequency, and Spatio-Temporal Domains.* Wiley, 2013.

[4] N. Chiras, C. Evans, and D. Rees, "Nonlinear gas turbine modeling using narmax structures," *IEEE Tans. Instrum. Meas.*, vol. 50, pp. 893–898, 2001.

[5] O. Boaghe, S. Billings, L. Li, P. Fleming, and J. Liu, "Time and frequency domain identification and analysis of a gas turbine engine," *Control engineering practice*, vol. 10, pp. 1347–1356, 2002.

[6] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *Journal of the Royal Statistical Society, Series B*, vol. 67, pp. 301–320, 2005.

[7] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society, Series B*, vol. 58, pp. 267–288, 1994.

[8] C. Novara, "Sparse identification of nonlinear functions and parametric set membership optimal analysis." in *in proc. American control conference, San Francisco, USA*, 2011.

[9] M. Bonin, V. Seghezza, and L. Piroddi, "Narx model selection based on simulation error minimisation and lasso," *IET control Theory and Applications*, vol. 4(7), pp. 1157–1168, 2010.

[10] R. De Keyser, *Model based Predictive Control for Linear Systems*. Invited chapter in UNESCO EoLSS. Oxford (6.43.16.1), 2003.

[11] C. Sprouse and C. Depcik, "Review of organic Rankine cycles for internal combustion engine exhaust waste heat recovery," *Applied Thermal Engineering*, vol. 51, no. 1-2, pp. 711–722, mar 2013.

[12] A. Desideri, B. Dechesne, J. Wronski, M. Van Den Broek, S. Gusev, V. Lemort, and S. Quoilin, "Comparison of moving boundary and finite-volume heat exchanger models in the modelica language," *submitted to energies*, pp. 1–17, 2016.

[13] S. Quoilin, R. Aumann, A. Grill, A. Schuster, and V. Lemort., "Dynamic modeling and optimal control strategy for waste heat recovery organic rankine cycles," *Applied Energy*, vol. Vol. 88, pp. 2183–2190, 2011.

[14] A. Hernandez, A. Desideri, C. Ionescu, S. Quoilin, V. Lemort, and R. De Keyser., "Experimental study of predictive control strategies for optimal operation of organic rankine cycle systems." in *Proceedings of the European Control Conference (ECC15)*, Linz, Austria, 2015.

[15] J. Peralez, P. Tona, O. Lepreux, A. Sciarretta, L. Voise, P. Dufour, and M. Nadri., "Improving the control performance of an organic rankine cycle system for waste heat recovery from a heavy-duty diesel engine using a model-based approach," in *In IEEE Conference on Decision and Control, Florence, Italy*, 2013.

[16] V. Grelet, P. Dufour, M. Nadri, V. Lemort, and T. Reichel, "Explicit multi-model predictive control of a waste heat rankine based system for heavy duty trucks," in *In IEEE Conference on Decision and Control, Osaka, Japan*, 2015.

[17] R. De Keyser and A. Hernandez, "Evaluation of the NEPSAC nonlinear predictive control on a thermal process." in *Proceedings of the 13th European Control Conference (ECC 2014)*, 2014.

[18] S. Quoilin, A. Desideri, J. Wronski, I. Bell, and V. Lemort., "Thermocycle: A modelica library for the simulation of thermodynamic systems," in *Proceedings of the 10th International Modelica Conference*, Lund, Sweden, 2014.

[19] V. Lemort, A. Zoughaib, and S. Quoilin., "Comparison of control strategies for waste heat recovery organic rankine cycle systems," *In Sustainable Thermal Energy Management in the Process Industries International Conference (SusTEM2011)*, 2011.

[20] L. Ljung, *System identification: theory for the user*. Prentice-Hall, 2007.