Utilization of Excitation Signal Harmonics for Control of Nonlinear Systems

Kasper Vinther, Henrik Rasmussen, Roozbeh Izadi-Zamanabadi and Jakob Stoustrup

Abstract—Many model based control methods exist in the literature. Producing a sufficient system model can be cumbersome and a new non-model based method for control of nonlinear systems with input/output maps exhibiting sigmoid function properties is therefore proposed. The method utilizes an excitation signal together with Fourier analysis to generate a feedback signal and simulations have shown that different system gains and time constants does not change the global equilibrium/operating point. An evaporator in a refrigeration system was used as example in the simulations, however, it is anticipated that the method is applicable in a wide variety of systems satisfying the sigmoid function properties.

I. INTRODUCTION

Most physical systems or processes are inherently nonlinear in nature, which makes design and tuning of controllers difficult, especially if the controller is expected to work in a large operation area, thus experiencing the full effect of the nonlinearity. Additionally, most systems are also time varying to some degree and change with unpredictable operating conditions and disturbances. This makes controller design even more difficult and often result in poorly tuned controllers. Possible ways to control these systems are with gain scheduling for time varying systems, see e.g. [1], [2], or adaptive backstepping control, see e.g. [3], [4].

Producing a sufficient nonlinear model for control purposes can be cumbersome, especially for time varying systems. Furthermore, a suitable reference operating point is often not known beforehand. A good alternative is therefore to look at non-model based methods such as extremum-seeking control, if it is desirable to drive the process output towards an extremum. If the input/output map does not have an extremum, but is instead "S" shaped exhibiting sigmoid function properties, then it is possible to use slope-seeking control. This is just a generalization of extremum-seeking control, where the reference slope is not zero. There are many examples of extremum- and slope-seeking control applied to practical control problems, see e.g. [5], [6], [7].

In this paper we look at a non-model based approach for control of time varying nonlinear systems, where the input/output map exhibits sigmoid function properties. A problem with slope-seeking control is that it can be hard to identify a reference slope, since the slope is often dependent on systems parameters that change with operating conditions. Another problem is that a badly chosen reference slope can make the system unstable, since a reference slope will be mirrored around the middle of the "S" shaped curve. If for example one chooses the slope at the middle of the curve as reference, then the slope will decrease in both directions leading to instability. An example of a system where the middle of the curve could be a desirable operating point is given in Section II.

We instead propose a novel control method called harmonic control for such types of systems. A Fourier analysis is conducted in this method to identify the distortion of an excitation signal as it passes the system nonlinearity. Taking the cross product between the first and second harmonic of the excitation signal gives an error signal which is negative and positive depending on which side of the middle of the "S" shaped curve the current operating point is located. In other words, qualitative knowledge about the system nonlinearity is used together with Fourier analysis of an excitation signal to generate an error signal for feedback purposes. The amplitude of the excitation signal can furthermore be adapted to better suite the current gain in the system, giving a controllable amount of excitation in the output. The proposed method has potential in any system having a nonlinear "S" shaped input/output map and it does not require any reference set point nor does it have the instability problem experienced with slope-seeking control. Replacing sensors with qualitative knowledge in the control loop also has the potential of reducing the commissioning costs and in some systems it might also be difficult or even impossible to measure certain states.

The structure of this paper is as follows. An example of a nonlinear system is first given in Section II, which will be used for simulation purposes. Then, the concept of harmonic control is given in Section III, which entails a presentation of the control strategy, how the error and control signal is generated, and adaption of the excitation signal amplitude. The harmonic control is then applied to the nonlinear system in simulation and compared with slope-seeking control and the results are presented in Section IV. Finally, conclusions are drawn in Section V.

II. SYSTEM NONLINEARITY

The nonlinear system given as example in this paper is an evaporator in a refrigeration system. Refrigeration systems normally operate by continuous vaporization and compression of refrigerant. This process is maintained by a valve, an evaporator, a compressor and a condenser, and this setup remains to a considerable extent the same. The
details of the vapor compression type refrigeration process are not given here, but can be found in e.g. [8].

Refrigeration systems are typically controlled by decentralized control loops and evaporator superheat is controlled in one of these loops by regulating the Opening Degree (OD) of a valve, see e.g. [9], [10], [11]. Superheating the refrigerant beyond the evaporation temperature is important, since no superheat means that two-phase refrigerant will enter the compressor and increase the power consumption and wear. Therefore, the valve flow must be kept at a level, where all refrigerant is evaporated before it reaches the compressor. However, there should be as much two-phase refrigerant in the evaporator as possible, to increase the heat transfer and thus optimize the refrigeration process.

The level of superheat in the evaporator is traditionally determined using at least a temperature sensor at the evaporator outlet and a pressure sensor, which can be used to calculate an evaporation temperature. The response from input to output in the evaporator is in general very nonlinear, making the output consist mainly of an offset and the first and second harmonic of the excitation signal.

A test was conducted on a refrigeration system with a water tank and heater as load on the evaporator. The system has an Electronic Expansion Valve (EEV) with controllable OD and controllable compressor and condenser fans. Sensors furthermore measure temperatures, pressures and flow and the system is monitored and controlled using the XPC toolbox for Simulink. In the test the OD of the EEV was slowly increased while the evaporator outlet temperature was measured (the compressor speed and condenser pressure was kept constant during this test). The result of this test is presented in Fig. 1. The evaporator outlet temperature $T_o$ for this refrigeration system follows an "S" shaped curve, and this system nonlinearity is then approximated using the inverse trigonometric function $atan$ shown in the bottom graph in Fig. 1. The mathematical expression is

$$ y = \left( -atan\left( \frac{4\pi (u - 50)}{50} \right) + \frac{\pi}{2} \right) 10, \quad (1) $$

where $y$ is the output (evaporator outlet temperature $T_o$) and $u$ is the input (OD). The OD is limited to a value between 0-100% open, where OD = 50% is located in the middle of the "S" shaped curve. This point in the "S" shaped curve also represents a close to optimal setpoint for the outlet temperature, as the evaporator is nearly filled (low superheat). A lower superheat is hard to maintain and does not provide that much safety margin.

The approximation in Fig. 1 is shown again in Fig. 2 where the system input has been overlayed with a sine excitation signal in three different operating points. It is clear from the output of the system that the sine excitation has been distorted. This is due to the nonlinearity in the system. In point (1) and (3) the sine is distorted more to one side than the other and in point (2) the distortion is equally large in both directions. Looking at the power spectrum density made over 80 periods of the excitation signal, it is observable that the output consist mainly of an offset and the first and second harmonic of the excitation signal.

A system time constant $T_{sys}$ and delay $T_d$ is furthermore added to the approximation of the output described in (1),

![Fig. 1](image1.png)

**Fig. 1.** The top graph shows the evaporation temperature $T_e$, outlet temperature $T_o$ and ambient water temperature $T_w$ for an evaporator during a sweep in the $OD$ of the EEV. The bottom graph is an approximation of the nonlinearity seen in the outlet temperature.

![Fig. 2](image2.png)

**Fig. 2.** The top graph shows the nonlinear input/output map with three operating points marked as (1), (2) and (3). The system is excited with a sine wave in each of these points and the steady state output is shown in the middle graphs along with a power spectrum density analysis in each operating point (bottom graphs).
given in the Laplace domain as
\[ G(s) = \frac{1}{1 + sT_{sys} e^{-sT_d}}. \] (2)
The value of \( T_{sys} \) and \( T_d \) can be chosen almost arbitrarily and in this simulation example they are set to 25 and 10 seconds, respectively. Note that (1) and (2) combined gives a Hammerstein model structure.

The system nonlinearity used in this paper is taken from an test refrigeration system, with valve OD as input and evaporator outlet temperature \( T_e \) as output. However, it is important to remember that the method applies in general for systems, with nonlinearities that can be approximated with sigmoid functions.

### III. HARMONIC CONTROL

This section introduces the concept of harmonic control. Various situations are analyzed through simulation and indicative conclusions are drawn based on the observed behavior of the excited system.

#### A. Control Strategy

The distortion seen in Fig. 2 can be used to generate an error signal. This distortion is measurable through Fourier analysis and requires a continuous excitation of the system, e.g., using a sine wave. The error signal \( \xi_n \) can then be used to drive the nonlinear system to the state with the highest gain (the middle of the "S" shaped nonlinearity), e.g. using Proportional-Integral (PI) control. Additionally, the Fourier analysis can give an estimate of the amplitude of the excitation in the output, which provides a way to adapt the amplitude of the excitation signal \( \alpha_e \) at the input, in order not to disturb the system more than what is necessary. The proposed control strategy called harmonic control is depicted in Fig. 3. In case there is saturation on the input it can help to add anti-windup to the integral part of the PI control. Furthermore, using the harmonics to generate the error signal acts as a powerful way of filtering the measurement. However, an adequate time scale separation between the system dynamics, the excitation frequency (Fourier analysis), and the PI controller is required, which is also the case with extremum- and slope-seeking control, see [5]. The next two subsections will address error signal generation and adaption of the excitation signal amplitude.

#### B. Error Signal Generation

The discrete Fourier series \( F(t) \) of a periodic function \( f(t) \), sampled at time \( t_n \) with sampling time \( T_{ex} \) (N is the number of samples in one period \( T_{ex} \)), is given as
\[ F(t) = a_0 + \sum_{p=1}^{M} \left[ a_p \cos \left( \frac{2\pi}{T_{ex}} p n \right) + b_p \sin \left( \frac{2\pi}{T_{ex}} p n \right) \right] \] (3)

\[ a_0 = \frac{1}{N} \sum_{n=1}^{N} f(t_n) \]
\[ a_p = \frac{2}{N} \sum_{n=1}^{N} f(t_n) \cos \left( \frac{2\pi}{N} p n \right), \quad p = 1, \ldots, \frac{N}{2} - 1 \]
\[ b_p = \frac{2}{N} \sum_{n=1}^{N} f(t_n) \sin \left( \frac{2\pi}{N} p n \right), \quad p = 1, \ldots, \frac{N}{2} - 1, \]

where \( a_p \) and \( b_p \) are the Fourier coefficients for each of the \( M \) harmonics denoted by \( p \). If taking \( N \) samples it is possible to determine \( M = N/2 \) harmonics. Mathematical study of the overlapping waves is called harmonic analysis and the Fourier coefficients of the first and second harmonic can be used to generate a scalar error signal \( \xi \) using the cross product of vectors in \( \mathbb{R}^2 \) given as
\[ \xi = \begin{bmatrix} a_1 \\ b_1 \end{bmatrix} \times \begin{bmatrix} a_2 \\ b_2 \end{bmatrix} = a_1 b_2 - a_2 b_1. \] (4)
The cross product between the first and second harmonic is also defined as
\[ \begin{bmatrix} a_1 \\ b_1 \end{bmatrix} \times \begin{bmatrix} a_2 \\ b_2 \end{bmatrix} = |H_1| |H_2| \sin(\phi) \] (5)
\[ |H_1| = \sqrt{a_1^2 + b_1^2}, \quad |H_2| = \sqrt{a_2^2 + b_2^2}, \]
where \( |H_1| \) and \( |H_2| \) are the amplitudes of the first and second harmonic respectively and \( \phi \) is the angle from the first harmonic to the second harmonic. The cross product is a normal vector and it is positive if the operating point is located to the right of the middle of the "S" shaped nonlinearity and negative to the other side. If there is enough time separation between system dynamics and excitation, then the two vectors in the cross product will be almost perpendicular. Equation (4) therefore provides a usable error signal for feedback. However, the error signal is still nonlinear, which can be corrected by making a normalization with respect to the cubed amplitude of the first harmonic.
\[ \xi = \frac{a_1 b_2 - a_2 b_1}{\sqrt{a_1^2 + b_1^2}} a_1 b_2 - a_2 b_1 = \frac{a_1 b_2 - a_2 b_1}{(a_1^2 + b_1^2)^{1.5}}. \] (6)
The normalized error signal \( \xi_n \) is calculated at different OD values in Fig. 4, 5, 6 and 7 under different situations to visualize the effects on the error signal. Note that the Fourier coefficients \( a_1, b_1, a_2, b_2 \) and \( b_2 \) are determined online in the implementation using time invariant linear FIR filters with the same sample time as the rest of the system and a ring buffer of size \( N \) in accordance with (3).

---

![Fig. 3. Block diagram of proposed closed loop harmonic control.](image-url)
Fig. 4 shows the error signal using three different system time constants. The normalized error remains linear in all cases. A change in the system delay $T_d$ has the same effect, however, an increase gives a decrease in the slope of the normalized error. Fig. 4 also indicates that the generated error can give a globally stable controller with only one equilibrium at the point of zero mean curvature. This is under the assumption that the nonlinearity fulfill the properties of a sigmoid function and that the succeeding controller is sufficiently slow compared to the excitation of the system.

Fig. 5 shows the error signal with three different system gains (the total gain in the system is changed $\pm 25\%$). Again the normalized error signal remains linear in all cases and the closed loop system will converge to the same point. This is not achievable with slope-seeking control, since the reference slope will have to be changed.

Fig. 4 and 5 also indicates that the slope of the normalized error signal increases when the time constants increases and decreases when the delay and gain in the system increases. Therefore, the controller gain should be chosen based on the worst case values of the system time constant, delay and gain, assuming that the system can be approximated with the dynamics described in (2) including a parameter varying nonlinear gain.

The system time constant and delay also determines the period of the excitation signal as this period should be slow enough to make the excitation visible in the output and at the same time as fast as possible to speed up the controller. Fig. 6 shows the error signal using three different excitation signal periods $T_{ex}$. Using $T_{ex} = 45$ makes the closed loop system loose global stability at low $OD$ values. Simulations have shown that the excitation signal period should at least be twice the combined system time constant and delay. Any higher period time gives approximately the same normalized error, which means that it does not affect the controller gain.

Fig. 7 finally shows the error signal using three different excitation signal amplitudes $a_e$. Changing the amplitude between 2% to 4% in $OD$ does not change the normalized error signal and the amplitude can therefore be adapted without compromising the controller. However, the amplitude should still be within reasonable bounds and large enough to overcome any noise and quantizations in the measurement.

One problem with Fourier analysis is that the Fourier series defined in (3) only converge to the output if the system is periodic. This is not the case when the system is operated in closed loop, especially during startup, unless the controller is tuned so slow that the output looks periodic. Fig. 8 shows an open loop simulation, where the $OD$ is kept constant and then slowly increased giving a periodic and an aperiodic part. The power spectrum density is calculated for each part and the aperiodic part has more low frequency content. However, reconstructing the signal using a Fourier series is
not required, since we only use the two first harmonics to calculate an error signal and these harmonics are still present.

C. Adaption of Excitation Signal Amplitude

The gain in the nonlinear system is largest at the equilibrium point. In order to be able to generate enough excitation of the output at low gains, while still maintaining an acceptable level of excitation when the gain is highest, it is often necessary to adapt the excitation amplitude. The harmonic analysis used to calculate the error signal can also be used for adaption. The amplitude of the first harmonic $|H_1|$, calculated as in (5), is a good approximation of the amplitude of the excitation at the output. Keeping this amplitude close to a reference can be achieved by changing the input excitation signal amplitude. However, it is undesirable to change the amplitude instantly in closed loop, since this will result in unstable behavior. The MIT rule is therefore used to adapt the amplitude slowly and it is defined as (see e.g. [12]):

$$J = \frac{1}{2} e^2$$

$$\frac{d\theta}{dt} = -\gamma \frac{\partial J}{\partial \theta} = -\gamma e \frac{\partial e}{\partial \theta}$$

where $J$ is an objective function to be minimized, $e$ is the error, $\theta$ is the adjustable parameter to be adapted and $\gamma$ is the adaption gain. The MIT rule can be interpreted as a gradient method for minimizing the error and in the case of adapting the amplitude $|H_1|$ we have

$$\theta = |H_1|$$

$$e_{|H_1|} = |H_1|_{ref} - |H_1|$$

$$\frac{dc}{dt} = \gamma e_{|H_1|},$$

since the partial derivative of $e_{|H_1|}$ is equal to $-1$. The amplitude $|H_1|_{ref}$ is the desired reference amplitude of the first harmonic. Only the adaption gain $\gamma$ has to be chosen. In general a small $\gamma$ means slow convergence and a large $\gamma$ means fast convergence and possibly instability. However, it is hard to say in general how $\gamma$ influences time variant systems, but a general rule is that the adaption must be slower than the control loop. Another possibility would be to use the normalized MIT rule, which would lead to less sensitivity towards signal levels or one could use Lyapunov stability theory to adapt the amplitude $|H_1|$, and most likely obtain faster adaption and stability guarantees.

IV. SIMULATION RESULTS

The performance of harmonic and slope-seeking control is compared in this section. The reference slope in slope-seeking was set manually to a value giving a reference operating point just above point (2) and below (1), see Fig. 2. Note that (2) is an unstable point in slope-seeking control.

The proportional gain $K_{p,h}$, the integral time $T_i$ and integral reset time $T_r$ were tuned manually for the harmonic controller and set to 0.2, 2 and 2 respectively. This gives a good compromise between convergence speed and robustness. The system can maintain a stable limit cycle at a critical controller gain, which can be found approximatively using the describing function method. Due to space limitations this analysis will not be part of this paper, but will be considered in future publications. The slope-seeking control gain $K_{p,s}$ was also tuned manually and set to 0.04, which is small enough to give acceptable oscillations in the control signal. The oscillations appear after the demodulator and a low pass filter can help filter this out [5].

The $OD$ signal has a saturation limit of 0 to 100%, and this was lowered in the controller to between 5 to 95% to allow the excitation signal to get through uncut. The adaption gain $\gamma$ in the harmonic controller should be slower than the PI control and was set to 0.00003, with a reference amplitude $|H_1|_{ref}$ of 2. Furthermore, the excitation signal amplitude was set to 2 in the slope-seeking controller. Fig. 9 shows the performance of the two controllers in a simulation without measurement noise and disturbances. The slope-seeking controller reaches a settling boundary of 5% of the step size about twice as fast as the harmonic controller, however, the harmonic controller can go to 50 in control signal without becoming unstable. If the reference slope was changed just slightly up or if the gain in the slope-seeking controller was raised giving more pronounced oscillations, it would make the controller go to the saturation limit of 95. The excitation signal amplitude was adapted during the
The excitation signal amplitude is slowly adapted from 5 to approximately 1.7, which makes the amplitude of the first harmonic $|H_1|$ converge to the reference of 2. The upper limit is 5 due to the saturation limits.

Fig. 11 shows the performance of the two controllers in a simulation with measurement noise ($2\sigma = 0.14^\circ C$) and a disturbance in the system nonlinearity. The performance is similar to the performance without noise, however, the harmonic controller experiences small oscillations in the control signal when the gain in the system is low ($u_c$ between 25 to 37), but it does converge to the correct operating point. Quantization on the measurement of 0.1$^\circ C$ and on the valve $C$ and on the valve $309$% was also added in the simulation shown in Fig. 11 to simulate realistic conditions. This quantization does not disturb the controllers noticeably.

The disturbance was made as an increase in the total gain in the system of 25%, see Fig. 5, introduced with a time constant of 10 seconds, 15000 seconds into the simulation. This corresponds to a decrease in the flow in the refrigeration system and causes the control signal in the slope-seeking controller to decrease as the reference slope now corresponds to a different operating point. If the disturbance was made as a decrease in the gain, then the slope-seeking controller would have become unstable. In comparison the harmonic controller does not deviate from the operating point.

The promising simulation results have lead to the implementation of the control method on two different test refrigeration system in our laboratory and the results are presented in [13].

V. CONCLUSION

A new non-model based method for control of nonlinear systems with input/output map exhibiting sigmoid function properties have been proposed. The method is called harmonic control and utilize an excitation signal together with Fourier analysis to obtain qualitative knowledge about where on the nonlinearity the closed loop system is located.

It was discovered that the excitation signal amplitude and period has little to no effect on the linear error signal as long as the period is about twice as large as the system time constant and delay. The amplitude of the signal can therefore be adapted to limit the unavoidable oscillations in the output due to the excitation. The control method was also simulated using different system gains and time constants, which did not change the global equilibrium/operating point. This was not the case with slope-seeking control. Furthermore, slope-seeking has stability problems if the reference slope is not chosen carefully and in general it can be difficult to obtain a suitable reference slope.

An evaporator in a refrigeration system was used as example in the simulations. However, it is anticipated that the method in general is applicable when the point of zero mean curvature of the system nonlinearity is close to the desired operating point and when the system input/output map satisfies the sigmoid function properties.