# Quantifying directed nonlinear coupling between brain and muscle activity:
## a NARX model-based approach

Nina van der Helm

# Quantifying directed nonlinear coupling between brain and muscle activity: a NARX model-based approach

by

N. A. van der Helm

to obtain the degree of **Master of Science**

at the Delft University of Technology,

to be defended publicly on Monday May 7, 2018 at 15:00 AM.

| | | |
|---|---|---|
| Student number: | 4167805 | |
| Project duration: | November 1, 2017 – May 5, 2018 | |
| Thesis committee: | Dr. ir. A. C. Schouten, | TU Delft, Supervisor |
| | Ir. K. E. Rodriguez Hernandez, | TU Delft, Supervisor |
| | Dr. F. M. Vos, | TU Delft, External Committee Member |
| | Prof. dr. ir. J. W. van. Wingerden, | TU Delft, External Committee Member |

An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

TU Delft

## Thank you

After nearly seven fantastic years of studying in Delft, I am ready to graduate and conquer the world as an engineer. Since this idea makes me feel a bit nostalgic, I would like to take the opportunity to write a, slightly dramatic, thank word.

To start, I want to thank Alfred and Karen for their support during the final year of my master. Alfred, you have the quality to make difficult problems seem simple. I often went to our meetings full of questions and doubts and returned with a clear plan of action. That was great. Karen, thanks for all the feedback on my writing, helping me through the struggles at the beginning of my literature study and, maybe most important, for the motivating emails.

Furthermore, I am very grateful for the warm support of my friends and family. I definitely believe in the healing power of a beer after a long, frustrating day of studying and you were there to drink it with me. There are two people I'd like to thank in specific. Kees, thank you for putting things in perspective (*"Stel je niet aan!"*), comforting me when I was stressed (*cookies!*) and helping me with some time-management issues. I probably have to listen to you more often. Finally, I want to thank Simon for the our conversations about math, his optimism about the ending of my project (*think: Coldplay - Fix you*) and, of course, for being my rock during the final days of finishing my thesis.

Nina

# Contents

# Quantifying directed nonlinear coupling between brain and muscle activity: a NARX model-based approach

**Nina A. van der Helm** **Ir. Karen Rodriguez Hernandez**
**Dr. ir. Alfred Schouten**

April 30, 2018

**Abstract**

**Relevance:** To enhance our understanding of motor impairments (e.g. post-stroke or due to Parkinson's Disease), objective measures for communication in the nervous system are required. By applying system identification techniques to oscillations in brain and muscle activity, we can objectively quantify the coupling between these areas.

**Gap:** Unfortunately, none of the existing techniques combine the ability to assess nonlinear behaviour and to detect causality in a closed-loop, which is necessary to fully characterize communication in the nervous system.

**Methods:** In this study, a new connectivity measure, the Nonlinear Directed Transfer Function (NDTF) is introduced. The NDTF is derived by mapping a nonlinear autoregressive model (i.e. NARX model) to the frequency domain, and provides an approximation of the linear and nonlinear causal influences on the output spectrum.

**Results:** The NDTF was validated using simulated data of a bidirectional, nonlinear system. Additionally, NDTF was applied to simultaneously recorded EEG-EMG of a wrist flexion task. For the experimental data, the NDTF results were dominated by linear interactions.

**Conclusions:** The NDTF has proven advantages above existing connectivity measures. However, it is sensitive to changes in the sampling frequency and segmentation, making interpretation difficult. The mainly linear interaction found in the EEG-EMG data implies limited sensory feedback, since the ascending pathways are known to act nonlinear.

## I. Introduction

Synchronization of neural oscillations is key to communication in the central nervous system. Neural populations oscillate together, and these oscillations can be transferred to spatially distant groups of neurons [1], [2]. During muscle contraction, activity in the contralateral sensorimotor cortex is coupled with activity in the muscles. These synchronization patterns encode information transfer between brain and muscles. By applying system identification techniques, such as coherence, to brain and muscle oscillations (i.e. corticomuscular coherence, CMC), it is possible to quantify corticomuscular coupling [3], [4], [5].

The importance of corticomuscular coupling becomes clear when looking at individuals with motor impairments. Stroke survivors with poor motor recovery show a significantly lower CMC than healthy individuals [6]. Furthermore, individuals with Parkinson's Disease (PD) off medication show CMC in a lower frequency range than controls and PD patients on medication [7], [8]. By quantifying and studying corticomuscular coupling in healthy individuals, we gain insight in the pathophysiology behind neurological movement disorders, like stroke or PD. In the long run, these insights might improve diagnosis and treatment.

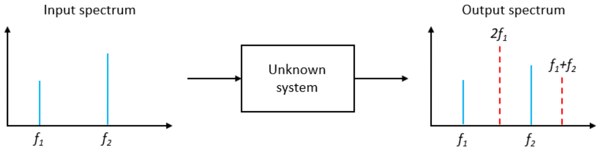



**Figure 1:** *(Adapted from Yang et al. [15]) The linear components of a system result in isofrequency coupling at $f_1$ and $f_2$ **(blue)**. The nonlinear part of the system results in cross-frequency coupling at harmonic ($2f_1$) and intermodulation ($f_1 + f_2$) frequencies **(red)**.*

Unfortunately, identifying corticomuscular coupling is challenging. Communication during motor control is bidirectional (i.e. closed-loop): the brain transmits motor commands to muscles and the muscle spindles and Golgi Tendon Organs (GTO's) send sensory feedback in return [9], [10]. Additionally, the nervous system acts highly nonlinear. Vlaar et al. quantified the nonlinear contribution to sensory evoked potentials (SSEPs), and found that nearly 80 percent of the response was caused by nonlinear processes [11]. The nonlinear behavior likely originates from the muscle spindles which are known to introduce second-order nonlinearities [12].

Although CMC is the most commonly used connectivity measure, it has two shortcomings: it is not able to describe causality within a loop [9] and it is a strictly linear measure [13]. Thus, to fully assess corticomuscular coupling, an alternative connectivity metric is needed that combines the ability to detect causality in a bidirectional system with the ability to describe both linear and non-linear interactions.

Yang et al. developed multi-spectral coherence (MSC) and multi-spectral phase coherence (MSPC) to cope with nonlinear coupling [14], [15]. Linear interactions are characterized by isofrequency coupling (i.e. coupling at the same frequency). In a nonlinear system (e.g. $y(t) = x^2(t)$) cross-frequency coupling occurs (Figure 1). MSPC and MSC are extensions of linear coherence and phase coherence that can capture cross-frequency coupling too. However, they are still not able to fully detect causality between signals.

Campfens et al. proposed to apply mechanical perturbations to the joint to deal with the closed-loop structure [16]. Since the perturbation signal is not part of the loop, coupling between the perturbation signal and brain activity mainly describes the dynamics of the ascending pathways. Choosing smart periodic perturbations can drastically improve the signal-to-noise ratio because it becomes possible to average over epochs of one perturbation period. However, this approach has three clear disadvantages. Firstly, the coupling can only be quantified on the perturbed frequencies and combinations or harmonics of those frequencies, making quantification dependent on the chosen perturbation signal. Furthermore, the perturbation signal pushes the nervous system in a certain artificial state. Therefore, it is hard to generalize the results to daily life activities. Lastly, the perturbation introduces an independent source for both the EEG and EMG signal, which makes simultaneously studying the ascending and descending pathways difficult [17].

Directed measures, such as the Directed Transfer Function (DTF) [18] and Partial Directed Coherence (PDC) [19], rely on linear autoregressive models instead of perturbations to separate ascending and descending information flow. This is a promising approach as it requires no perturbations to distinguish direction of information flow.

Only recently, the model-based directed approach has been extended to nonlinear interactions by using a nonlinear autoregressive (NARX) model [20] instead of its linear brother, the ARX model. He et al. derived a nonlinear variant of the PDC based on the NARX model, the NPDC, and applied it to EEG data to study corticocortical coupling [21]. The NPDC quantifies the (non)linear coupling for each output frequency. But, in an attempt to mimic the matrix formulation of the linear PDC, the nonlinear coupling at output frequency $f_i$ was divided by the input spectrum at $f_i$. This is incorrect: for nonlinear interactions, the output at $f_i$ is independent of the input at $f_i$ (Figure 1). Therefore, until today the *holy grail*, a method that can correctly separate direction and detect nonlinearities, was not found.

The aim of this study was to develop a new nonlinear connectivity measure that would meet both criteria. Based on the NARX-model, the Nonlinear Directed Transfer Function (NDTF) was derived, which is a measure for the input-output connectivity at each output frequency per order of nonlinearity. In contrast to the NPDC, the nonlinear part of the NDTF at $f_i$ is independent of the input at $f_i$. Simulated data was used to demonstrate the robustness of NDTF as a first step.

Subsequently, NDTF was used to quantify and compare sensory feedback during a wrist flexion position task and a wrist flexion force task as a proof of concept. The NDTF was determined for EEG and EMG data that was simultaneously recorded during both tasks. Based on previous studies, we assumed that the position task required more position feedback from the muscle spindles [22], [23], which show second-order nonlinearities [12]. Therefore, we expected a stronger second-order nonlinear coupling in the ascending pathways for the position task compared to the force task.

The remainder of the paper is outlined as follows. The methodology is divided in three sections. The theoretical derivation of the NDTF is explained in Section II. Subsequently, in Section III and Section IV, the application on simulated data and experimental data is described. Afterwards, both the simulation results and experimental results are presented in Section V. Interpretation of these results is described in the Discussion (Section VI). Finally, the most important outcomes of this study are summarized in the Conclusion (Section VII).

## II. Nonlinear Directed Transfer Function (NDTF)

In this section the Nonlinear Directed Transfer Function (NDTF) is derived. NDTF is a model-based connectivity measure, meaning that it quantifies information flow in two steps:

1. a parametric model is fitted to the data to describe the system, and
2. information flow is quantified by evaluating the fitted model.

The first two subsections describe the modeling procedure step-by-step (Subsection II.A-B). Eventually, the quantification step that results in the NDTF is explained in the last subsection (Subsection II.C).

### II.A. NARX model for nonlinear bidirectional systems

Modeling the human nervous system is challenging for two main reasons: information flows in two directions (i.e. cortex $\rightleftarrows$ muscles), and the interactions are both linear and nonlinear. This subsection explains how the NARX modeling structure addresses both problems by fitting a time-domain model to the data.

#### *a. Causality in a bidirectional system*

Model-based connectivity measures rely on the concept of Granger causality to detect direction of information flow in a closed-loop system. Granger basically stated that if signal $x_1$ causes signal $x_2$, $x_2$ can be (partially) predicted based on signal $x_1$ [24]. Vice versa, if it is possible to model signal $x_2$ using signal $x_1$, there must be a causal influence from $x_1$ on $x_2$.

For example, let $x_1$ and $x_2$ be two signals with discrete time recordings at $t = 1, 2, 3, ..., N$. The system linking signal $x_1$ and $x_2$ with can be described with a linear ARX model (i.e. AutoRegressive model with eXogenous input):

$$x_1(t) = \sum_{k=1}^{q} a_{11,k} x_1(t-k) + \sum_{k=1}^{p} a_{12,k} x_2(t-k) + e_{x_1}(t) \quad (1)$$

$$x_2(t) = \sum_{k=1}^{q} a_{21,k} x_1(t-k) + \sum_{k=1}^{p} a_{22,k} x_2(t-k) + e_{x_2}(t) \quad (2)$$

where $p$ and $q$ are the maximum time lags (i.e. the memory) of signals $x_1$ and $x_2$, $e_x(t)$ and $e_{x_2}(t)$ are the model prediction errors and $a_{11}$, $a_{12}$, $a_{21}$ and $a_{22}$ are the parameters fitted to minimize $e_{x_1}(t)$ and $e_{x_2}(t)$. The influence of signal $x_2$ on signal $x_1$ is reflected in parameter $a_{12}$. The other direction of information flow, signal $x_1$ influencing signal $x_2$, is represented by parameter $a_{21}$. Hence, the ARX model structure separated pathways $x_1 \to x_2$ and $x_2 \to x_1$.

#### *b. Nonlinear terms*

To quantify both linear and nonlinear coupling, nonlinear terms should be added to the model. Billings et al. extended the linear ARX structure with nonlinear terms, resulting in the NARX model [20]. Most common is the polynomial NARX model. Besides past input and output variables, the polynomial NARX model contains combinations of variables, representing nonlinearities in the system. The structure of a second-order nonlinear NARX model is as follows [25]:

$$\begin{aligned} x_2(k) = \theta_0 + & \underbrace{\sum_{i_1=1}^{n} \theta_{i_1} x_{i_1}(k)}_{1^{\text{st}} \text{ order terms}} \\ & + \underbrace{\sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \theta_{i_1 i_2} x_{i_1}(k) x_{i_2}(k)}_{2^{\text{nd}} \text{ order terms}} + e(k) \end{aligned} \quad (3)$$

with

$$x_{i_m} = \begin{cases} x_2(k-m) & 1 \leq m \leq n_{x_2} \\ x_1(k-(m-n_{x_2})) & n_{x_2}+1 \leq m \leq n_{x_2}+n_{x_1} \end{cases} \quad (4)$$

where $n$ is the total time lag $n = n_{x_1} + n_{x_2}$; $\theta_{i_m}$ and $\theta_{i_m i_m}$ contain the first- and second-order model parameters, respectively; $x_{i_m}$ contains lagged input and output terms as described by equation (4) and $e(k)$ is the prediction error. $\theta$ in this equation is equivalent to $a$ in equations (1) and (2).

*c. Model term selection*

ARX models are generally full models: they contain all possible terms up to a specified time lag. The impact of each term is controlled by the fitted parameters. Parameters of spurious terms will approach zero, which essentially eliminates these terms from the model.

In case of the NARX model, the number of terms increases exponentially with increasing time lags, since combinations of all terms are also taken into account. Parameter estimation of that many terms becomes computationally challenging. Therefore, a sparse model is more suitable: only meaningful terms are added to the model before parameter estimation. Model term selection was performed using the Forward Regression Orthogonal Least Squares algorithm based on the error-reduction-ratio (ERR) [26]. An elaborate explanation of this algorithm and the subsequent parameter estimation can be found in Appendix A.

To model both pathways in a closed-loop, *two* polynomial NARX model should be fitted separately (i.e. similar to equation (1) and (2), one for each direction).

## II.B. NARX models in the frequency domain

Although neural communication can be studied in the time domain (e.g. using event-related potentials), most studies focus on describing communication in the frequency domain. Fortunately, the NARX models identified in Subsection II.A can be mapped to the frequency domain.

A linear transfer function is simply expressed as $X_2(f) = H(f)X_1(f)$, where $H(f)$ is the transfer function. Linear interactions in the frequency domain are characterized by isofrequency coupling: input at frequency $f_1$ results in output at frequency $f_1$ (Figure 1). For nonlinear processes, input at frequency $f_1$ results in output at different frequencies. Here, we can distinguish between harmonic responses (e.g. $f_1 \rightarrow 2f_1$) and intermodulations (e.g. $f_1, f_2 \rightarrow f_1 + f_2$).

Therefore, the second-order transfer function

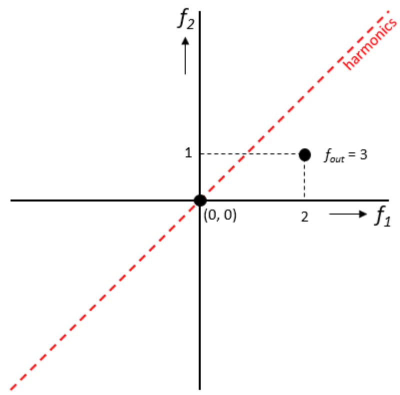



**Figure 2:** *The two-dimensional frequency space of a second order system. For each combination of frequencies, there is a complex number in $H_2(f_1, f_2)$ that defines the transfer from the input frequencies to the output frequency. The output frequency is the sum of the input frequencies. The red dashed line represents harmonic coupling, where $f_1 = f_2$ and thus $f_{out} = 2f_1$.*

$H_2(f_1, f_2)$ depends on two input frequencies and can be visualized as a 2D-surface. The output frequency at each point of the surface is the sum of its input frequencies (Figure 2).

Billings et al. developed a recursive method to compute $H_1(f_1)$ and $H_2(f_1, f_2)$ using the estimated parameters and terms from the NARX time-domain model [27], which is summarized in Appendix B.

## II.C. From NARX model to NDTF

While the implications of a first-order transfer function are easy to interpret, this becomes more difficult for the second order. Furthermore, based on the separate transfer functions $H_1$ and $H_2$, it is not possible to see the combined effect of the first and second order terms. The last step, turning the frequency-domain NARX model into the new connectivity measure, NDTF, is meant to make interpretation and visualization of the modeling results easier.

$H_{1,x_1 \leftarrow x_2}$ and $H_{2,x_1 \leftarrow x_2}$ both reflect a part of the coupling from $X_2(f)$ to $X_1(f)$. By calculating the result of $H_2$ for each **output** frequency, the transfer is reduced to a 1D function.

The recorded physiological signals (i.e. EEG and EMG) contain a broad spectrum of frequencies. Different combinations of these frequencies will certainly overlap in output frequency (e.g. $f_1 = 3, f_2 = 1 \rightarrow f_{out} = 4$ and $f_1 = 2, f_2 = 2 \rightarrow f_{out} = 4$). The influence of each of these frequency combinations depends on the input spectrum at $f_1$ and $f_2$ and on the coupling strength defined in $H_2(f_1, f_2)$ [25]. To determine the total influence on $f_{out}$, all frequency combinations that add up to $f_{out}$ can be summed. The NDTF is simply this sum of contributions:

$$\text{NDTF}_{1,x_1 \leftarrow x_2}(f) = |H_1(f)X_2(f)| \quad (5)$$

$$\text{NDTF}_{2,x_1 \leftarrow x_2}(f) = \left| \sum_{f=f_1+f_2} H_2(f_1, f_2)X_2(f_1)X_2(f_2) \right|, \quad (6)$$

where $f$ is the output frequency, $H_1$ and $H_2$ are the transfer functions for the first and second order nonlinearities, and $X_2(f)$ is the input spectrum. The second-order NDTF is a discrete approximation, since only the terms that are available in vector $x_1$ are taken into account. The discretization inevitably causes a loss of combinations and, therefore, a lower NDTF. We assumed that the loss would be similar across conditions and across the frequency spectrum. A visual example of the consequences of discretization on $\text{NDTF}_2$ is given in Appendix C.

Finally, the contributions of the first and second order can be summed, such that the combined effect can be studied.

$$\text{NDTF} = \text{NDTF}_1 + \text{NDTF}_2 \quad (7)$$

If the system is linear, or if we are only interested in the linear part of a system, the second term of the summation can be left out.

## III. Simulations

To prove that our NARX-based framework is valid, simulations were performed on a system that globally mimicked the dynamics of the corticomuscular loop. The simulated system (i.e. Data Generating System (DGS)) that was used is described in Subsection III.A. Subsequently, the data analysis protocol is outlined in Subsection III.B. The metrics used to quantify the performance of NDTF as a connectivity measure are discussed in Subsection III.C. To identify and address potential problems in the use of NDTF for EEG/EMG data, some simulation experiments were performed. These are described in Subsection III.D.

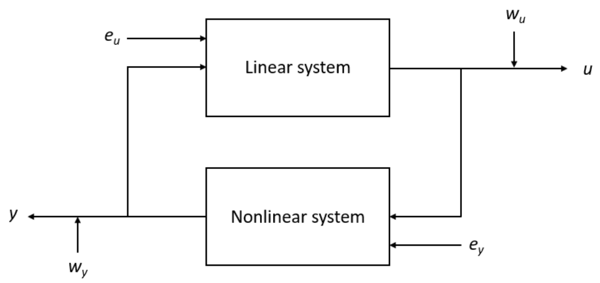



**Figure 3:** *Block scheme of the simulated system as described in He et al. [21]. It is a bidirectional system with output signals u and y and white noise sources $e_u$ and $e_y$ that excite the system. The transfer $u \leftarrow y$ is linear; the transfer $y \leftarrow u$ is nonlinear. The default simulation settings for the system are provided in Table 1*

### III.A. Data Generating Systems

There were two main requirements for the Data Generating System (DGS):

1. the system had to be bidirectional, and
2. the system had to contain nonlinearities.

He et al. developed a numeric example of a system that meets both requirements (Figure 3). The system has two outputs, $y$ and $u$. The processes that connect $u$ and $y$ can be described with the following two difference equations:

$$\begin{aligned} y(k) &= 0.5y(k-1) - 0.3y(k-2) + 0.1u(k-2) \\ &\quad + 0.4u(k-1)u(k-2) + e_y(k) + w_y(k) \end{aligned} \quad (8)$$

$$\begin{aligned} u(k) &= 0.3u(k-1) - u(k-2) \\ &\quad - 0.1y(k-2) + e_u(k) + w_u(k) \end{aligned} \quad (9)$$

where $u$ and $y$ are the outputs of the system; $e_y$ and $e_u$ are zero-mean Gaussian white noise sources ($var(e) = 0.01$) that excite the system, and $k$ is the discrete time-step. Signal $u$ depends linearly on the past of signal $y$. On the other hand, signal $y$ depends non-linearly on the past of signal $u$ due to a second-order nonlinear input term ($0.4u(k-1)u(k-2)$).

### III.B. Data analysis

For the every condition in the simulation experiments, the system was simulated 10 times. The data analysis was performed separately on each of these repetitions to be able to estimate the standard deviation over repetitions afterwards.

| Parameter | Default settings |
|---|---|
| recording time $T$ | 1000 [s] |
| sampling frequency $f_s$ | 20 [Hz] |
| variance $e_u$ | 0.01 |
| variance $e_y$ | 0.01 |
| variance $w_u$ | 0 |
| variance $w_y$ | 0 |
| number of segments $L$ | 100 |

**Table 1:** *Default settings for the DGS depicted in Figure 3. 1000 [s] recording time at 20 Hz results in 20000 samples. These are split up in 100 segments of 2000 samples. There is no measurement noise added in the default simulations. Therefore, the variance of $w_y$ and $w_u$ is zero.*

First, the acceptable residual prediction error was determined based on a grid search across different thresholds. Multiple second-order nonlinear NARX models (equation (3)) were fitted with varying thresholds for the residual prediction error. The residual prediction error was expressed as the error-to-signal ratio (ESR) [25],

$$ESR = 1 - \sum_{i=1}^{M} ERR_i, \qquad (10)$$

where $ERR_i$ is the error-reduction-ratio for model term $i$ and $M$ is total number of terms added to the model. The calculation of the ERR is explained in further detail in Appendix A. The correct value for ESR depends on the amount of noise in the system, which is often unknown. The different models were evaluated based on their fit (i.e. VAF, see Subsection III.C) on a separate part of the data, that was not used during fitting. The ESR that resulted in the best model, was picked. This resulted in $ESR = 0.03$ for linear system $u \leftarrow y$ and in $ESR = 0.3$ for non-linear system $y \leftarrow u$.

The maximum time lags ($p$ and $q$ in equation (3)) were higher than what would be necessary based on our knowledge of the system ($n_u = 10$ and $n_y = 10$). However, this creates an opportunity to demonstrate that the NARX model converges to correct terms, even when there is an extensive library of other terms available.

### III.C. Performance metrics

To characterize the algorithms performance, four performance metrics were used.

1. *Variance-accounted-for (VAF)*
The VAF represents the variance explained by the time-domain model expressed as a percentage of the total variance in the output signal. It was used to asses the fit in the time domain. A high VAF corresponds to a low prediction error. The generated data was divided in two parts. 80 percent of the data was used to fit the model. The other twenty percent was used to calculate the VAF. Therefore, the VAF does not increase when the model fits noise in the training data. The VAF can be calculated as follows:
$$\text{VAF} = \left(1 - \frac{\text{var}(\boldsymbol{y} - \hat{\boldsymbol{y}})}{\text{var}(\boldsymbol{y})}\right) \cdot 100\% \qquad (11)$$
where vector $\boldsymbol{y}$ is an epoch of the output signal in the validation data, and $\hat{\boldsymbol{y}}$ is the modeled signal for the same epoch.
2. *Quality metric (Q)*
The linear part of each pathway in the DGS acts as a narrow passband filter, which primarily passes components around 4.5 Hz (for a 20 Hz sampling frequency). Therefore, we expect the NDTF to find coupling at 4.5 Hz for both systems. Additionally, we expect a peak at the second-order harmonic: 9 Hz ($2 \times 4.5$ Hz) for the interaction $y \leftarrow u$. Quality metric Q is the ratio between output on frequencies where output is expected and output on frequencies where there is not. It is expressed as follows (adapted from Potgieter et al. [28]):
$$Q = \frac{\text{NDTF}_{\mu}^{f_{exp}}}{\text{NDTF}_{\mu}^{f_{unexp}}} \qquad (12)$$
The NDTF is averaged over the range of either expected or unexpected frequencies. The expected frequency range is 4-5 Hz for $u \leftarrow y$ and 4-5 and 8.5-9.5 Hz for $y \leftarrow u$; the unexpected frequency range includes all other frequencies. A high Q value means that the NDTF accurately describes the system in the frequency domain. Bear in mind that Q can only be calculated for simulated systems where the $f_{exp}$ and $f_{unexp}$ are known beforehand.
3. *Peak-value of NDTF*
The third performance measure is simply the peak-value of the NDTF in the 4-5 Hz and 8.5-9.5 Hz region.
4. *Standard error of the mean (SEM)*
The standard error of the mean reflects the accuracy of the estimated parameters. For correct

model terms in a time-invariant system, the parameters should be approximately constant for each sample, resulting in a low SEM. The SEM for each variable can be found on the diagonal of the covarance matrix $Cov_{\hat{\theta}}$, which is determined as follows [13]:

$$Cov_{\hat{\theta}} = \frac{1}{N}\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}\left(P^TP\right)^{-1}, \tag{13}$$

Here, $N$ is the total number of samples, vector $\boldsymbol{\epsilon}$ contains the prediction error for each sample, and each column in matrix $P$ contains the samples of one of the selected model terms

### III.D. Simulation experiments

The simulation experiments are used to evaluate how the following parameters influence the effectiveness of NDTF: 1) variations in data segmenting, 2) variations in coupling strength, and 3) additive measurment noise.

1. *Data segmenting*
   By default, the system is simulated for 1000 seconds at 20 Hz, resulting in $N = 20000$ samples. The data is then segmented into $L$ non-overlapping epochs, a part of which is used for fitting and a part of which is used for validation. Data segmenting influences the frequency resolution.

$$\Delta f = \frac{f_s}{N_{epoch}} = \frac{f_s \cdot L}{N} \tag{14}$$

   where $\Delta f$ represents the frequency resolution; $f_s$ is the sampling frequency; $N_{epoch}$ is the number of samples per segment; $L$ is the number of segments and $N$ is the total number of available segments. During this simulation experiment, analysis was additionally performed for $L = 50$, which resulted in segments of 4000 samples.

   We expected that a higher frequency resolution will result in a higher NDTF since the discrete approximation of the amplitude spectrum becomes better (i.e. more frequency combinations are taken into account). The loss-effect as described in Appendix C will, therefore, probably decrease. Additionally, longer segments will result in less leakage to neighboring frequencies.

2. *Coupling strength*
   A stronger causal coupling in one of the pathways should result in a higher NDTF. To test this, the coupling strength was altered in two ways, in separate simulations:
   - the causal gain $\theta_{y(t-2)}$ in the linear system $u \leftarrow y$ was increased from 0.1 to 0.15, and
   - the gain for the nonlinear term, $\theta_{u(t-1)u(t-2)}$, in system $y \leftarrow u$ was increased from 0.4 to 0.6.

   The expectation is that the higher gains will increase the peaks of the NDTF. The increase of $\theta_{y(t-2)}$ will likely affect the linear peak at 4.5 Hz in system $u \leftarrow y$. The increase of $\theta_{u(t-1)u(t-2)}$ will probably affect the second-order peak at 9 Hz, for system $y \leftarrow u$.

3. *Effect of measurement noise*
   In the default simulations, no measurement noise was added to the signal. However, physiological systems often show high levels of both process and measurement noise. To verify whether the algorithm still performs well in the presence of noise, in this experiment white Gaussian noise was added to the output signals.

   The addition of noise was regulated based on the noise-to-signal ratio (NSR):

$$NSR_y = \frac{\text{var}(\boldsymbol{w_y})}{\text{var}(\boldsymbol{y})} \times 100\% \tag{15}$$

   where $\boldsymbol{w_y}$ is the additive measurement noise and $\boldsymbol{y}$ is the actual signal. Simulations were performed for two situations: symmetric noise and asymmetric noise. In the symmetric case, the NSR was kept equal for signal $u$ and $y$. Simulations were run for $NSR = 0\%, 10\%, 20\%$ and $50\%$.

   In the asymmetric case, the same values for the NSR were used. However, the measurement noise was only added to signal $y$. We expected that output noise results in a larger effect on the NDTF than input noise.

## IV. Acquisition of EEG/EMG data

This section explains the methodology for collecting the EEG-EMG data.

### IV.A. Participants

Four participants volunteered in the study (1 female; age: 23 ± 1.4 years). All of the participants were self-reported right-handed and had no history of neurological diagnosis or injuries to the right arm. Each of them provided written informed consent prior to the start of the experiment. The experimental protocol was approved by the Human Research Ethics Committee (HREC) of the Delft University of Technology.

### IV.B. Experimental setup

An overview of the experimental setup is given in Figure 4. The setup was located in a sound-proof cabin with dimmed light, to avoid distraction or overstimulation of the participant. The participant was seated next to a haptic wrist manipulator (Wristalyzer, Moog Inc., the Netherlands). The right forearm was strapped to the arm rest of the manipulator such that the flexion angle of the elbow was approximately 90°. The height of the manipulator was altered for each subject to a position where both shoulders were relaxed. The arm rest and hand grip were positioned such that the pivot point of the wrist and the manipulator were aligned. The hand was strapped to the hand grip to avoid grasping. In front of the participants was a monitor that provided visual feedback during the tasks. Direct view of the right hand was shielded by a black canvas.

High-density EEG was measured using a 128-channel cap (WaveGuard cap, ANT Neuro) with Al/AgCl electrodes. Simultaneously, EMG was measured using bipolar Al/AgCl electrodes on the extensor and flexor carpi radialis muscle belly with a 1.5cm inter-electrode distance. The EEG signals, EMG signals and the exerted wrist torque and angle were recorded at 2048 Hz on a bio-signal amplifier (ReFa, TMSi, the Netherlands). The amplifier had a gain of 26.55 for physiological input channels and filtered the signals with a 552 Hz low-pass filter to prevent aliasing.

### IV.C. Experimental protocol

All participants performed different two tasks: a force task and a position task.

1. *Force task*
   The hand grip was fixed at 0° wrist flexion. For each trial, participants were instructed to exert a isotonic flexion torque (1 Nm) with their right wrist during 35 seconds. Beforehand, they had 10 seconds to build up and stabilize their wrist torque. Visual feedback was continuously provided on the monitor. The target torque corresponded to the arrow pointing upwards.
2. *Position task*
   The hand grip was unlocked and could be moved frictionless through the range of motion. For each trial, participants were asked to keep their wrist angle constant at a 0° angle during 35 seconds, while withstanding a 1 Nm torque that was independent from the angle. Prior to the hold period, they had 10 seconds to move to the 0° angle and to stabilize there. Visual feedback was again provided on the screen, where 0° corresponded to the arrow pointing up. To stimulate the use of sensory feedback, angular deviations were amplified on the screen, such that a 1° angular deviation corresponded to a 16° deviation of the arrow on the screen.

Two participants started with the force task; the other two started with the position task. For each task, the participant performed two training trials followed by ten actual trials. Between trials there was a 15 second break. In total, the experimental protocal took approximately 30 minutes. The entire experiment took approximately 2.5 hours, including informed consent, instructions, setting up the EEG and EMG and cleaning afterwards.

### IV.D. Data preprocessing

For each trial, the first second of data was discarded to remove transient behavior. Subsequently, for both the EEG and EMG channels, the mean was subtracted and the data was bandpass filtered (1-200 Hz) with a 4$^{th}$ order Butterworth filter. Additionally, the 50 Hz noise from the power line and its harmonics at 100 and 150 Hz were removed with a notch filter. The data was then cut into 2 second non-overlapping epochs, resulting in 170 epochs per participant per condition.

Bad channels (i.e. with high variance or with an impedance above 20 $k\Omega$) and epochs containing clear deviations (e.g. due to eye blinks) were removed from the EEG data based on visual inspection. The EEG was transformed with a spherical Laplacian derivation to filter the data spatially. The electrode that showed the strongest coherence with the EMG signal from the flexor carpi radialis was selected to calculate the NDTF. The coherence was averaged between 15 and 30 Hz, which is the frequency-range where CMC is usually highest [4], [29]. The EMG signal from the flexor carpi radialis was included as the second input for the NARX model. Finally, both signals were downsampled from 2048 to 512 Hz before fitting the NARX model.

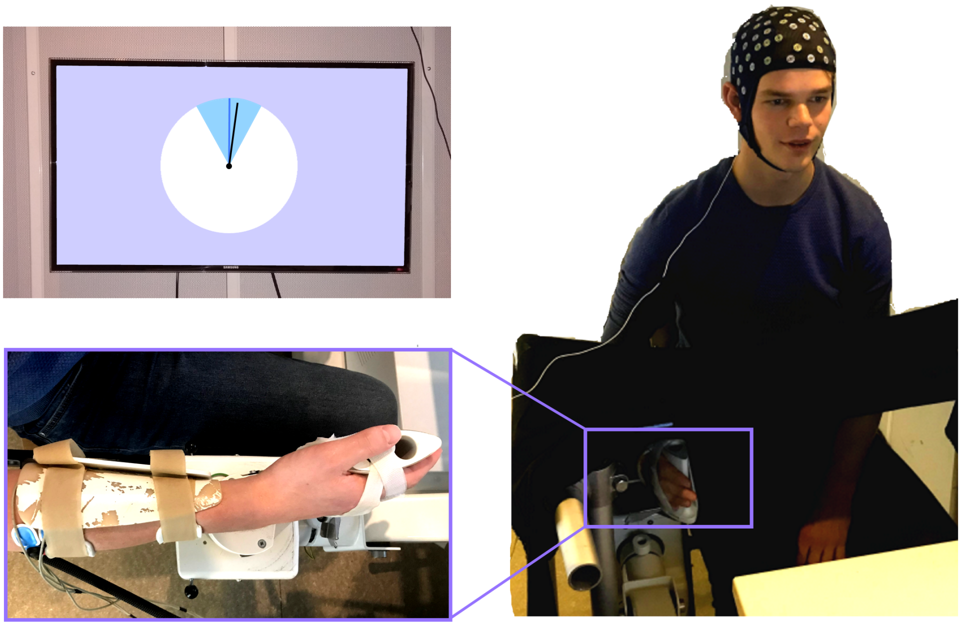


**Figure 4:** *[**Right**] Overview of the participant in the setup. Participants is seated next to the haptic manipulator. Direct view of the right wrist is shielded. [**Left, top**] Monitor in front of the participant, providing them with visual feedback during the tasks. The target position/torque is the line straight upwards; the actual position/torque is shown by the dark blue arrow. [**Left, bottom**] The right fore-arm strapped to the haptic device. The axis of the wrist and the manipulator are aligned. EMG electrodes are placed at the flexor carpi radialis and the extensor carpi radialis (only one of the electrodes is visible in the picture).*

### IV.E. Data analysis

Per participant, four different NARX models were fitted (i.e. two tasks, two directions). Eighty percent of the data was used to fit the NARX model, the other 20 percent was used to evaluate the models performance.

Similarly to the simulation analysis, models were built for different values of ESR (equation (10)). The model was computed for $ESR = [0.1, 0.15, 0.2, \text{ and } 0.25]$. These values were selected based on a grid search across a broader range of ESR values. The different models were compared based on their VAF (equation (11)) and the best model was selected.

The maximum output lag $n_{out}$ is set at 10 timesteps, which corresponds to approximately 20 ms at a 512 Hz sampling frequency. The minimum input time lag was set at 8 steps (i.e. approximately 15 ms) to account for the conduction delay between motor cortex and muscles. The maximum input time lag was set at 18 timesteps (i.e. approximately 35 ms). In total, 10 input and 10 output steps were taken into account for both pathways.

To obtain a good discrete approximation of the NDTF, a high frequency-resolution is beneficial. The frequency resolution depends on the length of each segment (equation (14)). Therefore, 111 segments (i.e. the minimum number of epochs available after preprocessing) were concatenated after fitting. A Hann window was applied to each segment before attaching them to avoid leakage. The input spectrum for the NDTF was obtained by taking the Fourier transform of the long, concatenated time-series.

A low frequency-resolution has the advantage that the bandwidth of the resulting spectrum is small and, thus, easy to interpret. To smoothen the high-resolution NDTF, a moving-average filter with a 20 step-window was applied as a final step.

## V. Results

This section contains the results of the simulation study (Subsection V.A) and the results based on the EEG-EMG data (Subsection V.B).

### V.A. Simulation results

This subsection contains the results of the simulation experiments done to demonstrate the feasibility of NDTF as a connectivity measure. First, the simulated output spectra $|U|$ and $|Y|$ are discussed. Secondly, the resulting NDTF for the default settings is presented. Lastly, the outcomes of the simulation experiments as described in Subsection III.D are listed.

#### *a. Simulated output spectra*

The amplitude spectra for signal $u$ and $y$ are shown at the top row of Figure 5. Output spectrum $|U|$ shows a peak around 4.5 Hz. $|Y|$ shows a peak at the same frequency, but also shows a second peak around 9 Hz, due to the second-order nonlinearity in the system. The magnitude of the first peak differs across both paths: 0.06 for system $u \leftarrow y$ versus 0.01 for system $y \leftarrow u$. Therefore, the process noise is relatively larger for system $y \leftarrow u$, which explains the noisy output spectrum (Figure 5).

#### *b. NDTF for default settings*

The results based on the default settings (Table 1) are shown in the 2$^{\text{nd}}$ row of Figure 5. In the figure, the first and second order NDTF are separated, to show where the peaks originate from. The linear system $u \leftarrow y$ shows no peaks in the second-order NDTF but a strong peak around 4.5 Hz in the first-order NDTF. For the non-linear system $y \leftarrow u$ there are two peaks visible: one around 4.5 Hz based on the first-order NDTF and one around 9 Hz for the second order NDTF. The NARX algorithm selected the correct model terms for all ten repetitions. The parameters for spurious terms approached zero, making their contribution limited. Parameters for correct model terms were correctly and consistently estimated. They can be found in Table 2 and 3 for system $u \leftarrow y$ and $y \leftarrow u$, respectively. The SEM plots on the bottom row of Figure 5 show clear differences between correct and incorrect terms when it comes to consistent parameter estimation. Quality measure Q, indicating the ratio between desired and undesired output, is substantially higher for $u \leftarrow y$ than for $y \leftarrow u$: 38.9 and 4.9, respectively. However, these two values can't be compared directly, since they are calculated differently (Subsection III.C) and describe different systems. The predictive capability of the NARX model in the time domain can be assessed based on the VAF. The VAF was 96.2 % and 57.72 % for $u \leftarrow y$ and $y \leftarrow u$, respectively.

#### *c. Simulation experiments*

Three simulation experiments were performed to study the effect of data segmenting, coupling strength and measurement noise on NDTF. The performance measures for these experiments are summarized in Table 2 and Table 3.

1. *Data segmenting*
   The number of segments was decreased from 100 to 50, with 4000 samples per segment instead of 2000. The resulting NDTF for both pathways is depicted in Figure 5. The longer segments resulted in a higher frequency resolution ($\Delta f = 0.005$ instead of $\Delta f = 0.01$). This has three consequences.

   Firstly, the lower number of segments results in worse performance of FROLS algorithm. The algorithm overlooked the term $y(t-2)$ in the linear system 4 out of 10 times. Without this term, there is no causal influence from $y$ to $u$, resulting in a flat NDTF. Secondly, $Q$ clearly increases for system $u \leftarrow y$ ($Q = 79.5$ instead of 38.9). For system $y \leftarrow u$, the quality remains nearly equal ($Q = 4.17$ instead of $Q = 4.89$. Thirdly, the peak value of the first peak (4.5 Hz) increases substantially for the linear system $u \leftarrow y$ (0.16 instead of 0.06).

2. *Coupling strength*
   For the second experiment, the gain of the causal term in the linear system and the causal squared term in the nonlinear system were, separately, multiplied by a factor 1.5. The new parameters were estimated correctly for both systems. For linear system $u \leftarrow y$ the gain of term $y(t-2)$ was increased from -0.1 to -0.15. Its parameter was estimated at $\theta_{y(t-2)} = -0.1533 \pm 0.0095$. For nonlinear system $y \leftarrow u$, the gain of nonlinear term $u(t-1)u(t-2)$ was increased from 0.4 to 0.6. $\theta_{u(t-1)u(t-2)}$ was estimated to be $0.5998 \pm 0.0059$.

   The higher coupling strength resulted in higher peak values for the NDTF. The peak around 4.5 for linear system $u \leftarrow y$ increased from 0.0629 to 0.0946, which is a 50,3 % growth (i.e. very close to the 50% increase in coupling strength). For nonlinear system $y \leftarrow u$, the focus should be on the second-peak, caused by the squared term. This peak increases from 0.0015 to 0.0017, which is a 13,3 % increase.

3. *Measurement noise*
   Lastly, the effect of adding white Gaussian

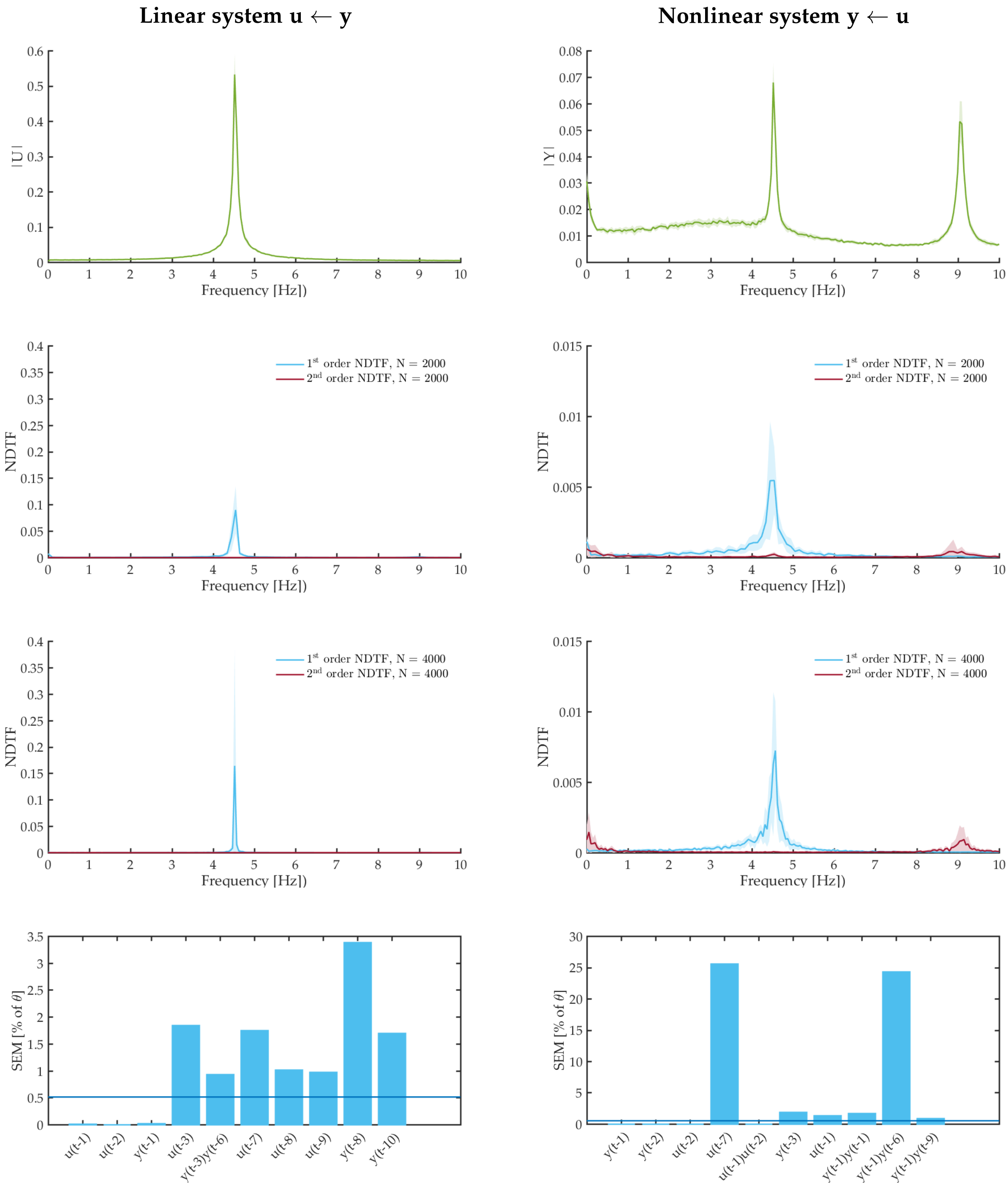



**Figure 5:** *Results of simulation and data analysis of the DGS (Subsection III.A). The figures on the left reflect the outcomes for the linear system $u \leftarrow y$; the figures on the right reflect nonlinear system $y \leftarrow u$.* ***(1st row)*** *Magnitude of the output spectra $|U|$ and $|Y|$ for the default settings (Table 1). The linear system (left) shows a clear peak value around 4.5 Hz. The influence of process noise $e_u$ is negligible. The nonlinear system (right) shows a peak at 4.5 Hz and a second peak at 9 Hz, which is caused by the second order non-linearity in the system. Process noise $e_y$ causes the blurring at the other frequencies.* ***(2nd row), (3rd row)*** *The 1st and 2nd order NDTF for segments of 2000 and 4000 samples, respectively. A higher number of samples results in a higher frequency resolution, which decreases leakage and increases the 2nd order NDTF since the discrete approximation becomes better (Appendix C).* ***(4th row)*** *Accuracy of the estimated parameters per model term for the default settings represented by the standard error of the mean (SEM). The SEM is expressed in percentage of the mean estimated $\theta$. The dark blue line indicates significance. All correct model terms are below the line (i.e. significant); no incorrect terms are significant. The 10 terms that are shown are the terms that were selected most often during the ten simulation repetitions.*

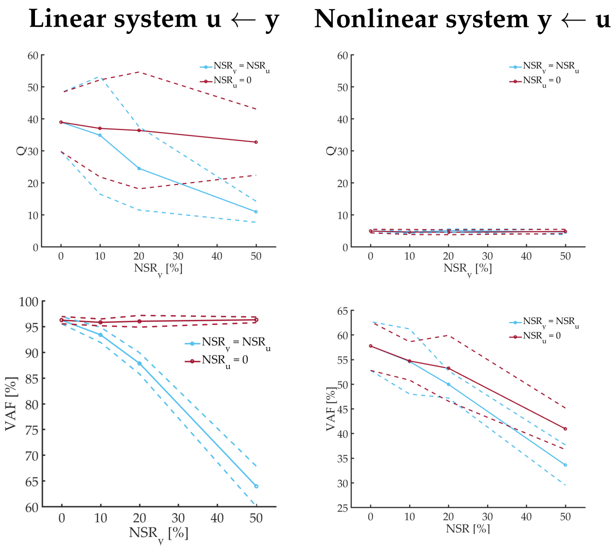



**Figure 6:** *Effect of white Gaussian noise on the model fit in the time domain, expressed by VAF* (11)*, and on the quality of the NDTF, expressed by Q* (12)*. On the x-axis is the noise-to-signal ratio for signal y as expressed in equation* (15)*, where NSR = 0 means noise-free. The blue line indicates 'symmetric noise' (i.e. equal NSR for signal u and signal y); the red line indicates 'asymmetric noise' (i.e. only additive noise on signal y).*

measurement noise to the output signals was investigated. There were two separate noise cases studied: symmetric noise and asymmetric noise. The influence of the noise cases on the VAF and Q for both pathways is depicted in Figure 6. The FROLS algorithm remained successful in selecting the correct model terms, even for the highest NSR. However, the parameters are underestimated in the presence of measurement noise (Table 2 and 3) and decrease with increasing NSR.

For linear system $u \leftarrow y$, the VAF decreases substantially in the symmetric noise case (from 96.2 % to 63.95%) and not at all in the asymmetric noise case (96.2 % versus 96.3 %). More important is the effect on Q and the peak value. Both Q and the peak value decrease substantially for both noise cases, although the decrease is larger for the symmetric case (Table 2).

For nonlinear system $y \leftarrow u$, the VAF decreases for both noise cases (default: VAF = 57.7 %, symmetric: VAF = 33.57 %, and asymmetric: VAF = 40.9 %). The quality of the measure Q, however, only decreases very minimally. There is a decrease visible in the peak values for both the first peak (0.0110 versus 0.0105 versus 0.0083) and the second peak (0.0015 versus 0.0011 versus 0.0007), where the symmetric noise case always results in the lowest peak.

## V.B. Experimental results

The results for the EEG-EMG data are divided in two parts: results of the data preprocessing and results of the data analysis (i.e NARX modeling and calculation of NDTF).

### a. *Preprocessed EEG-EMG data*

After removal of bad epochs based on visual inspection, at least 111 epochs were left per participant per task (143 epochs on average). The electrode with the strongest coherence was in most cases, as desired, located above the sensorimotor cortex (4x CP1, 2x C3, 1x C1 and 1x CP1). However, for subject s03 no clear coherence peak was located. Here, electrode CP1 was selected since this was the most commonly selected electrode. Topoplots of the corticomuscular coherence between 15-30 Hz are provided in Figure 7. The selected electrode is marked with a circle.

### b. *NDTF applied to EEG-EMG data*

The NDTF for all four participants in both directions and for both tasks is shown in Figure 8. There is no consistent difference between the force and the position task visible across participants. The subsequent paragraphs provide more detailed findings for the EEG-EMG-data.

In the time-domain, the predictive capability of the model can be described with the VAF (equation (11)). There was a substantially higher VAF found for the ascending pathways compared to the descending pathways (75.9 ± 5.6 versus 48.9 ± 5.3). This difference is similar for both the position task and the force task. Additionally, the descending and ascending pathways differed substantially in their peak values and average values for the NDTF. For the descending pathways (i.e. EMG ← EEG), these peaks lay between 0.002 and 0.05 (Figure 8). The ascending pathways, on the other hand, show substantially lower peak values between $10^{-}7$ and $10^{-}6$.

Figure 9 shows the NDTF for the first and second order combined, and the NDTF for only the first order (i.e. linear). The figure shows data for subject s01 on the force-task. The linear component of the NDTF clearly dominates the nonlinear component.

| | $\theta_{u(t-1)}$ | $\theta_{u(t-2)}$ | $\theta_{y(t-2)}$ | Peak-value | VAF | Q |
|---|---|---|---|---|---|---|
| **True system** | **0.3** | **-1** | **-0.1** | | | |
| Default settings | 0.3107 $\pm$0.0612 | -1.0035 $\pm$0.0261 | -0.0982 $\pm$0.0085 | 0.0629 $\pm$0.0202 | 96.2309 $\pm$0.7092 | 38.9278 $\pm$9.1791 |
| Longer segments | 0.3229 $\pm$0.0359 | -1.0040 $\pm$0.0162 | -0.1017 $\pm$0.0074 | 0.1642 $\pm$0.2193 | 96.7607 $\pm$0.4634 | 79.5557 $\pm$45.6505 |
| Symmetric noise | 0.0871 $\pm$0.0344 | -0.3462 $\pm$0.0296 | 0.0663 $\pm$0.0199 | 0.0159 $\pm$0.0086 | 63.9590 $\pm$3.9081 | 10.9252 $\pm$3.2818 |
| Asymmetric noise | 0.3197 $\pm$0.0836 | -1.0183 $\pm$0.0415 | -0.0639 $\pm$0.0067 | 0.0391 $\pm$0.0163 | 96.3022 $\pm$0.5521 | 32.6883 $\pm$10.3402 |
| **True system** | **0.3** | **-1** | **-0.15** | | | |
| Linear coupling $u \leftarrow y\ \uparrow$ | 0.3379 $\pm$0.0423 | -1.0214 $\pm$0.0197 | -0.1533 $\pm$0.0095 | 0.0946 $\pm$0.0620 | 95.2372 $\pm$0.5959 | 31.4040 $\pm$13.3016 |

**Table 2:** *Overview parameter estimates and performance metrics (Subsection III.C) for linear system $u \leftarrow y$ ($\mu \pm \sigma_d$). **(Default settings)**The default settings are provided in Table 1. The other systems all deviate from the default settings at one point. **(Longer segments)** Less segments with more samples per segments: $L = 50, N = 4000$ instead of $L = 1000, N = 2000$. **(Symmetric noise)** Additive measurement noise (w in Figure 3) on both output signal u and input signal y such that $NSR_u = NSR_y = 50\%$. **(Asymmetric noise)** Additive measurement noise solely on input signal y such that $NSR_y = 50\%$. **(Linear coupling $u \leftarrow y\ \uparrow$)** Increase in the gain of causal term $y(t-2)$, $\theta_{y(t-2)} = 0.15$ instead of* 0.1.

| | $\theta_{y(t-1)}$ | $\theta_{y(t-2)}$ | $\theta_{u(t-2)}$ | $\theta_{u(t-1)u(t-2)}$ | $1^{st}$ peak | $2^{nd}$ peak | VAF | Q |
|---|---|---|---|---|---|---|---|---|
| **True system** | **0.5** | **-0.3** | **0.1** | **0.4** | | | | |
| Default settings | 0.5113 $\pm$0.0334 | -0.3154 $\pm$0.0315 | 0.0980 $\pm$0.0052 | 0.3979 $\pm$0.0222 | 0.0110 $\pm$0.0037 | 0.0015 $\pm$0.0008 | 57.7296 $\pm$4.9476 | 4.8928 $\pm$0.5587 |
| Longer segments | 0.5185 $\pm$0.0670 | -0.2998 $\pm$0.0251 | 0.1041 $\pm$0.0089 | 0.4106 $\pm$0.0128 | 0.0092 $\pm$0.0037 | 0.0015 $\pm$0.0010 | 59.5046 $\pm$3.6592 | 4.1742 $\pm$0.5075 |
| Symmetric noise | 0.0985 $\pm$0.0659 | -0.0342 $\pm$0.0533 | 0.0307 $\pm$0.0031 | 0.0333 $\pm$0.0074 | 0.0083 $\pm$0.0042 | 0.0007 $\pm$0.0004 | 33.5750 $\pm$4.0813 | 4.6173 $\pm$0.8191 |
| Asymmetric noise | 0.2525 $\pm$0.0729 | -0.1143 $\pm$0.0600 | 0.1034 $\pm$0.0101 | 0.3711 $\pm$0.0534 | 0.0105 $\pm$0.0051 | 0.0011 $\pm$0.0008 | 40.9341 $\pm$4.2076 | 4.7755 $\pm$0.6623 |
| **True system** | **0.5** | **-0.3** | **0.1** | **0.6** | | | | |
| Nonlinear coupling $\uparrow$ $y \leftarrow u$ | 0.4998 $\pm$0.0090 | -0.3001 $\pm$0.0120 | 0.1000 $\pm$0.0030 | 0.5998 $\pm$0.0059 | 0.0093 $\pm$0.0028 | 0.0017 $\pm$0.0011 | 65.6716 $\pm$4.6950 | 4.0622 $\pm$0.5501 |

**Table 3:** *Overview parameter estimates and performance metrics (Subsection III.C) for nonlinear system $y \leftarrow u$ ($\mu \pm \sigma_d$). **(Default settings)**The default settings are provided in Table 1. The other systems all deviate from the default settings at one point. **(Longer segments)** Less segments with more samples per segments: $L = 50, N = 4000$ instead of $L = 1000, N =$* 2000. ***(Symmetric noise)** Additive measurement noise (w in Figure 3) on both input signal u and output signal y such that $NSR_u = NSR_y = 50\%$. **(Asymmetric noise)** Additive measurement noise solely on output signal y such that $NSR_y = 50\%$. **(Nonlinear coupling $y \leftarrow u\ \uparrow$)** Increase in the gain of causal term $u(t-1)u(t-2)$, $\theta_{u(t-1)u(t-2)} = 0.6$ instead of* 0.6.

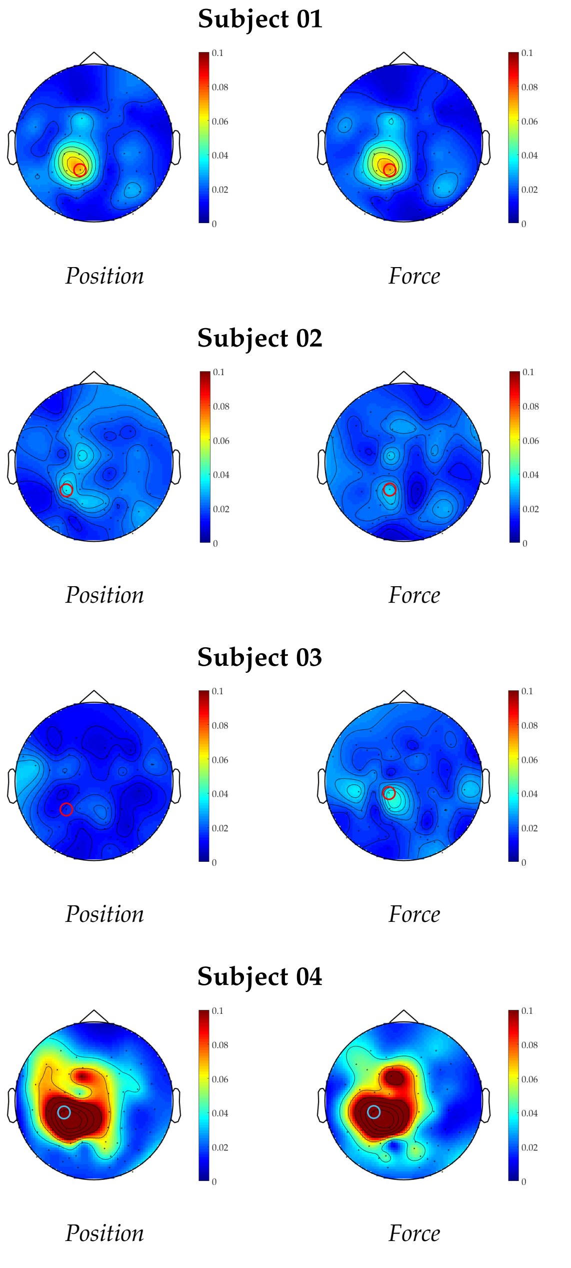



**Figure 7:** *Topoplots showing the coherence with the EMG signal for the carpi radialis flexor, averaged between 15 and 30 Hz. The electrode showing the strongest coherence was selected for further analysis. These electrodes are marked with a circle.*

The results for this participant were representative for all four participants during both tasks in terms of the ratio between linear and nonlinear contributions.

Lastly, to confirm our hypothesis, we focus on the second-order nonlinearities in the ascending pathways. The ascending pathways carry sensory feedback, and presumably contain second-order nonlinear dynamics. The $NDTF_2$ of the ascending pathways for each participant is plotted separately in Figure 10. Based on these four participants, no consistent difference can be observed in the ascending $NDTF_2$ between the force-task and the position-task.

## VI. Discussion

Many studies have shown that corticomuscular connectivity is both nonlinear [11], [15] and bidirectional [10], [9], [30]. Although nonlinear and bidirectional techniques exist, they are seldom combined to detect connectivity in neural systems. In this study, we introduced a new approach to quantify directed (non)linear corticomuscular coupling: the Nonlinear Directed Transfer Function (NDTF). The NDTF is derived from a NARX model. The NARX model structure has been used before and its relevance to describe a non-linear system has been proven on a great variety of (technical) systems [31], [32], [21]. However, the NARX model has not yet been used to characterize corticomuscular coupling. We propose a novel approach by using a discrete approximation of the causal output spectrum (NDTF) as connectivity measure.

The discussion starts with implications of this study on the field. In the subsequent two subsections (Subsection VI.B-C), the main findings of this study are listed and interpreted. First, the capability of NDTF as a connectivity measure was demonstrated using simulations. Subsequently, the feasibility of using NDTF on real EEG-EMG data was shown. The discussion addresses the findings based on these two phases separately. Afterwards, the limitations of this study are listed as methodological considerations (Subsection VI.D). Finally, the last subsection gives recommendations for future studies (Subsection VI.E).

### VI.A. Implications

The long-term goal of using system identification techniques to characterize corticomuscular coupling is to gain more insight in the impaired mechanisms

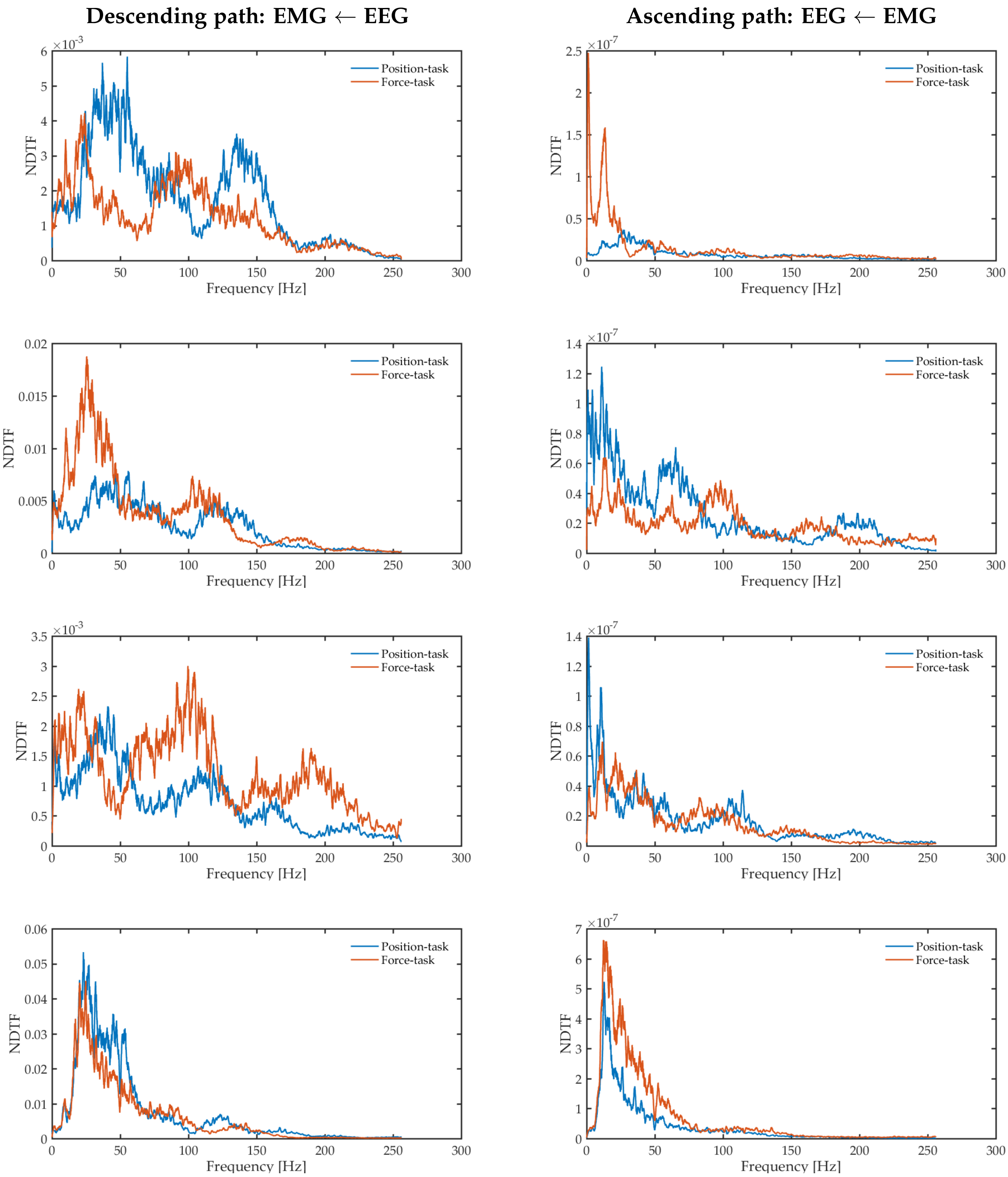



**Figure 8:** *The NDTF per participant per pathway for both tasks. On the left, the estimated NDTF for the descending path (i.e. EMG ← EEG) is shown; on the right, the NDTF for the ascending path is depicted. The shown NDTF is the sum of the first and second order NDTF. However, as shown in Figure 9,the NDTF is dominated by the linear contribution. There is no consistent difference between the force and position task. Lastly, notice the difference in scale between the acending and the descending pathways. The NDTF for the descending pathways is substantially higher that for the ascending pathways, in all participants.*

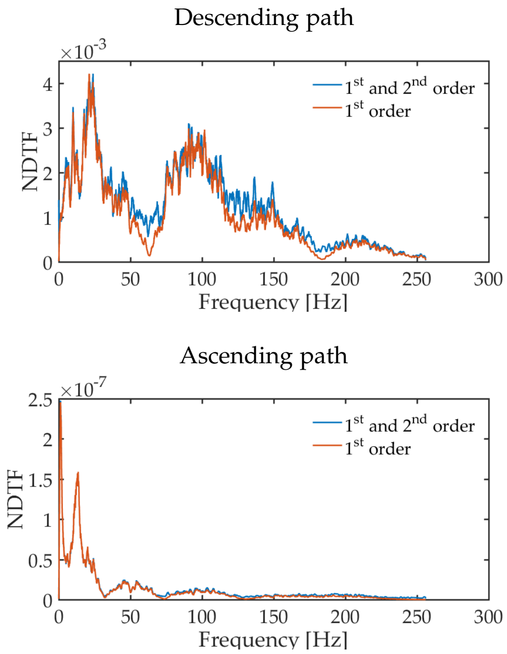



**Figure 9:** *First order NDTF, $NDTF_1$ (5), and total NDTF (i.e. sum of first and second order) shown separately for the force task of participant s01. The graph shows that the contribution of the second order is minimal compared to the first order, for both the ascending and descending pathways.*

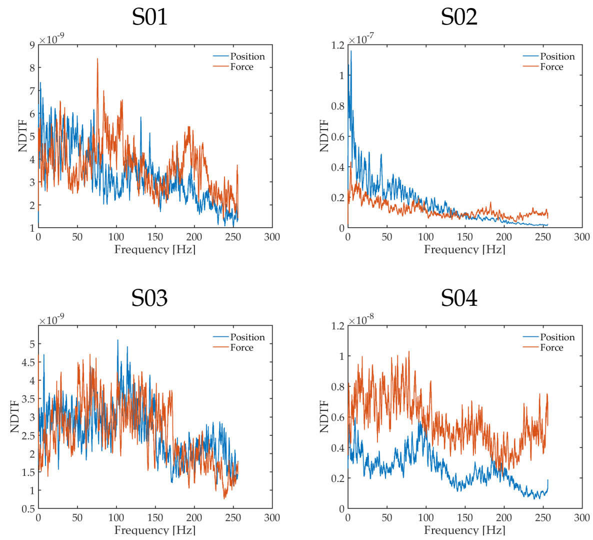



**Figure 10:** *The second-order NDTF, $NDTF_2$ (6), for the ascending pathways is plotted for each participant for both tasks. The ascending pathways carry sensory feedback. Since the muscle spindles, sensing position of the limb, are known to introduce second-order nonlinearities [12], a higher $NDTF_2$ in the ascending pathways was expected during the position task. This was only seen for participant s02. Participant s04 shows completely opposite behavior: a higher $NDTF_2$ for the force task.*

behind movement disorders. In the last decades, identification has focused on identifying the descending pathways [33], [34]. However, for many movement disorders it is essential to include the influence of the somatosensory system, i.e. the ascending sensory pathways. For example, the intactness of the sensory pathways is a good predictor of motor recovery in stroke survivors [35]. In individuals with PD, the resting tremor shows signs of an afferent input aside from the motor drive [36]. To study the ascending pathways, it is important to take nonlinear processes into account, since these pathways presumably behave highly nonlinear [11]. The NDTF is the first measure that can study nonlinear coupling in the ascending pathways, without using perturbations. Therefore, it can contribute substantially to our understanding of movement disorders.

### VI.B. NDTF on simulated data

Based on simulated data, we demonstrated that NDTF can detect the linear and nonlinear coupling for both pathways in a bidirectional system. The FROLS algorithm for model term selection proved to select the correct model structure [26], even in the presence of noise and with an extensive library of model terms to choose from. Due to the 1D mapping of the $2^{nd}$ order transfer function $H_2(f_1, f_2)$, the combined effect of the first and second order terms can be plotted in one coördinate system. NDTF could even be expanded to third or higher order nonlinearities. Therefore, NDTF is a useful addition to currently available measures to assess corticomuscular coupling.

Nevertheless, the NDTF as connectivity measure may present a couple of drawbacks. Firstly, the NDTF is sensitive to changes in the frequency resolution. A low frequency resolution results in leakage and worsens the discrete approximation of the second-order NDTF. A Hann window has proven useful in preventing leakage [37] and was, therefore, applied on the experimental data. The loss of information due to the discrete approximation, however, is harder to resolve (Appendix C). Only combinations of frequencies are taken into account that are represented in the discrete input spectrum. For example, $f_1 = 4$ Hz and $f_2 = 3$ Hz might result in an output at $f_{out} = f_1 + f_2 = 7$ Hz. But, input frequencies $f_1 = 3.9$ Hz and $f_2 = 3.1$ Hz might also have an effect at 7 Hz, although these frequencies might not be in the discrete frequency vector. So, by increasing the frequency resolution, more combinations of frequencies are added, and the approximation will become better. To compare the NDTF for two signals, it is important that the frequency resolution is equal, as it influences the magnitude of NDTF.

Secondly, while noise does not affect model term selection, it does influence the parameter estimation. Parameters are underestimated in the presence of measurement noise, leading to a lower NDTF. Parameter estimation in the symmetric case was more biased for both paths, which is reasonable considering that more noise was added in total. Moving towards real data, the effect of asymmetric noise should be studied more in-depth, since EEG contains significantly more noise than EMG. A next step could be to apply NDTF to simulated data where the total amount of noise is kept constant, but the distribution over signal $u$ and signal $y$ varies.

### VI.C. NDTF on EEG-EMG data

The feasibility of NDTF for real data was tested using EEG and EMG data recorded during two wrist flexion tasks: a force task and a position task. We expected a stronger ascending coupling (i.e. more sensory feedback) during the position task [22],[23]. Secondly, we expected to see second-order nonlinearities in the ascending coupling originating from the muscle spindles [12].

Based on the EEG-EMG data we can cautiously draw three conclusions: 1) the linear interactions were dominant in both pathways, 2) the descending NDTF was much higher than the ascending NDTF, and 3) the VAF is much higher for the ascending pathways than for the descending pathways.

Firstly, the NDTF for the EEG-EMG data for all four participants, shows that the linear part dominates the nonlinear part in both directions, while this was only expected for the descending pathways. Vlaar et al. [11] stated that nearly 80% of the sensory evoked potential was the consequence of nonlinear processes, originating mainly in the muscle spindles. This is supported by the widely accepted muscle-spindle model of Mileusnic [12] , which also contains second-order nonlinearities. Based on these studies, we expected a contribution of the second-order NDTF. There are two main explanations for the fact that a large second-order contribution is missing. First of all, the linear behavior could be explained by the fact that most other studies quantified nonlinearities using mechanical perturbations. Yang et al. studied nonlinear coupling based on a nonlinear extension of coherence, n:m coherence, without perturbations and found more nuanced results [38]. Although they did find some second-order nonlinear behavior, the largest part of the system was found to be linear. Mechanical perturbations push the nervous system in a certain, artificial state that might contain more nonlinear behavior.

Although the absence of perturbations probably plays a role, we pose that the attenuation of the nonlinear terms is actually inherent to the way $NDTF_2$ was derived. It comes down to the discrete approximation of the second-order NDTF. For linear systems, output at frequency $f_1$ is caused by input at frequency $f_1$ (Figure 1). The transport between these two is stored in the transfer function. Hence, multiplication of the transfer function with the input spectrum accurately describes the output spectrum. In case of nonlinear transfer with a continuous input spectrum, there is an infinite amount of combinations of frequencies that together result in the final output spectrum (Figure 2). Not all these combinations are taken into account. Therefore, the contribution of the second-order to a system is underestimated, while the contribution of the linear terms is not (Appendix C). This is supported by the simulation results where the second peak in the NDTF is much lower than the first, while this is not the case for the actual output spectrum (Figure 5). The approximation can be improved by increasing the frequency resolution and, thus, the number of combinations taken into account, but for a continuous spectrum it will always be inferior to the linear contribution.

Secondly, there was a clear difference between the magnitude of the NDTF found in the descending compared to the ascending pathways. The descending pathways showed NDTF peaks in the order of $10^{-3}$, while the ascending pathways stayed behind around $10^{-7}$. Previous studies that compared information flow based on linear ARX models, have found similar results [29], [39], [40]. They concluded that the descending pathways dominated motor control. We argue that the difference in magnitude is mainly explained by the difference in the variances of EEG and EMG. The muscle amplifies the neural input it receives from the motor neurons. The variance of the recorded EEG signals is approximately a factor $10^3$ smaller than the variance of the EMG signal. This results in high gains in the descending pathways and low gains in the ascending pathways, making comparing the two troublesome.

Lastly, the NARX models predictive capabilities, assessed with the VAF, were high for both the descending and the ascending pathways considering the high noise levels in EEG and EMG data. This can be explained by the large role of local dynamics. The VAF we calculate is based on the one-step-ahead (OSA) prediction, meaning that the error is reset after every step [25]. For a high sampling frequency (512 Hz), model term $y(t-1)$ becomes so close to $y(t)$ that the VAF, due to this one model term, becomes very high. However, this does not necessarily mean that the model structure is completely correct, just that the one-step-ahead prediction is close. Additionally, the model fit for the ascending pathways was much better than for the descending pathways ($VAF \approx 50\%$ and $VAF \approx 70\%$, respectively). The cause likely lies in the asymmetric distribution of noise in the signal. The EEG signal contain considerably more noise than the EMG signal. All that noise is amplified in the descending pathways, making the EMG signal hard to predict.

### VI.D. Methodological considerations

A few considerations for interpreting the current results are listed.

1. Simulations are performed on a discrete system, where the influence of all correct terms is picked in the same order of magnitude. However, for continuous, physical systems, the influence of each term depends on the sampling frequency. For physical systems at a high sampling rate, signal $y(t-1)$ is often close to $y(t)$ due to inertia. At the same time, the influence of $u$ decreases with decreasing step size, making it less likely for the causal terms to be selected by the algorithm. This is a well-known issue in model term selection [41]. A solution that has been used previously is to downsample the data, to increase the relative impact of the input terms.
2. The correct ESR depends on the noise level in the system, which is unknown. To approach the optimal value, multiple models with different ESR values were fitted and evaluated on a test data set. However, this was only a rough grid search so the accepted ESR value might still not be optimal. Even small differences in ESR can have a large impact on the eventual model.
3. It is not possible to validate whether the correct model terms are selected. The VAF only determines how well a model predicts an outcome but is no guarantee for a correct model structure. Since the frequency-domain NARX model is derived of the time-domain model, the right model structure is more important than the quality of the prediction. Some solutions for this limitation are discussed in the subsequent Subsection VI.E.
4. After initial parameter estimation, a linear noise model consisting of terms $e(t-1)$ and $e(t-2)$ is added to the selected model terms. A second linear-least-squares algorithm is applied to make sure that the parameters are not biased by past noise terms. However, only two linear terms are added. Noise cross-terms (e.g. $e(t-1)y(t-1)$) are not added and these terms might still bias the fitted parameters [20]. Noise enters the neuromuscular system at different points, not only as an addition to the output. Noise terms that enter the system earlier in the process will likely correlate with input or output terms going through that same process. Therefore, the parameters for these terms might be influenced if the noise model is incomplete.
5. During preprocessing of the EEG data, the electrode was selected based on the highest corticomuscular coherence. As coherence is a linear measure, this selection criterion might favor electrodes with a stronger linear component over electrodes with second-order nonlinearities. However, Vlaar et al. found similar electrodes while using the signal-to-noise ratio per electrode as selection criterium, which has no bias towards linear behavior [11]. So, the chances that this approached biased our results is limited.

### VI.E. Recommendations and future work

The recommendations are divided over four areas: 1) easy interpretation of the NDTF, 2) improvements in model-term selection, 3) validation of the model structure, and 4) improvements in data collection.

First, the NDTF is not an intuitive, easy-to-interpret measure. Many linear connectivity measures, such as coherence or PDC, vary between 0 and 1, making it relatively easy to interpret them and make comparisons between conditions, participants and even different studies. Such a range is achievable if a maximum value or an optimum of the measure is known, and that is in case of the NDTF a problem. He et al. tried to solve this problem by molding the nonlinear transfer function $H_n$ such that it would fit the linear matrix formulation that underlies PDC and DTF [21]. However, in this attempt, the nonlinear coupling at output frequency $f_1$ became dependent of the input spectrum at $f_1$. Since there is no isofrequency coupling for nonlinear interactions (Figure 1), this derivation was incorrect.

A solution might be to fit a continuous spectrum through the available data. A continuous spectrum has the advantage that no frequency combinations are lost, and therefore, the predicted outcome spectrum should approach the actual causal influence on the output spectrum. Billings already derived equation (6) for a continuous input spectrum [27], [42]. However, to bring this theory to practice, a method is required to convert discrete data to a continuous spectrum. If this continuous NDTF is normalized with respect to the measured output spectrum, the measure is approximately bound between 0 and 1.

Secondly, correct model term selection is key to a meaningful NDTF. In this study, the FROLS algorithm was used with the error-reduction-ratio of as selection algorithm, which is a computationally easy model term selection algorithm [26]. Many studies focused on improving the FROLS-ERR algorithm[41], [43]. Globally, the improvements can be divided into three topics: different or additional selection criteria, different stopping criteria and methods to remove spurious terms. An example is, for instance, to use mutual information as a selection criterion in addition to the error-reduction-ratio [41]. We propose that, to remove spurious terms, the standard error of the mean (SEM) values of the parameters might be used (Figure 5, bottom row). Terms that do not belong in the system, will presumably not fit to every data segment. Therefore, these incorrect terms can be recognized by a high SEM. A more elaborate selection algorithm could result in a better model fit and therefore, a more useful NDTF.

The third direction for future studies is the statistical validation of the NDTF. Statistical analysis of model-based measures like the NDTF is challenging. It is possible to prove three things: 1) how well the NARX model predicts the signal in the time domain based on the VAF, 2) how consistently the model parameters are estimated based on the SEM, and 3) whether the residual noise is white and independent of both the input and output signal with correlation tests [44] (not performed in this study due to time constraints). However, it is impossible to prove whether the correct model terms are selected for unknown systems. Furthermore, there is no confidence limit known for the NDTF.

In his book, Billings shows that the multiple-step-ahead prediction error (or simulation error) can indicate whether the correct model structure was chosen [25]. As mentioned before, we calculate the VAF based on the one-step-ahead (OSA) prediction. To accurately assess whether the model terms are chosen correctly, a multiple-step-ahead (MSA) prediction is the better choice. In this approach the error accumulates over steps, while for the OSA prediction the error is reset every step. Taking many steps ahead is not realistic for physiological data that contains high noise levels. However, calculating the VAF based on the, for example, five-steps-ahead prediction error, might give insight in whether the chosen model terms are correct.

A possible approach to estimate a confidence limit, might be random permutation (RP). This approach was previously used by Florin et al. [45] to assess the confidence limit for several linear model-based connectivity measures. Both the EEG and EMG data samples are randomly shuffled in the time domain, such that there is presumably no correlation between the two signals. The confidence limit can be estimated based on the NDTF for the non-correlated data.

The last area to focus on in future studies is improving the experimental protocol. The number of participants in this study was sufficient for a proof of concept. However, considering the large interindividual differences that have been found in similar studies [4], [30], a larger participant group is required to draw meaningful conclusions about the neuromuscular system.

A final note is that currently no individual component analysis (ICA) was performed on the data. One electrode was selected per participant to perform the data analysis. If an ICA is combined with a source localization algorithm, the resulting signal might contain less noise [46] and, therefore, result in a better fit for the NARX model.

## VII. Conclusion

The aim of this study was to introduce a new connectivity measure that could provide insight in linear and nonlinear processes in the human nervous system: the NDTF. The NDTF was demonstrated on simulated data and validated on real EEG-EMG data. This section summarizes the conclusion drawn in both phases of the study.

Based on the simulated data, the following conclusions were drawn:

- The NDTF can correctly identify both linear and nonlinear coupling in a bidirectional system and responds to changes in coupling strength.
- The magnitude of the NDTF increases with the frequency resolution of the input spectrum.
- The model term selection algorithm is robust to additive white noise. There is room for improvement in parameter estimation in the presence of noise.

Based on the EEG-EMG data, the following conclusions were drawn:

- Linear coupling seems dominant in both the ascending and the descending pathways for a constant force and constant position wrist flexion task. However, NDTF has a bias towards linear interactions, which could explain these results.
- The NARX model fit, assessed by the VAF, for the ascending pathways is consistently better than for the descending pathways.
- There was no consistent difference in NDTF across participants, between the wrist flexion force-task and position-task.

## Acknowledgements

I am grateful to Alfred Schouten and Karen Rodriguez Hernandez for their guidance during my master project, Joost Kordelaar and Mark van der Ruit for their assistance with data collection, Henri Boessenkool for his help with the experimental hardware and the ethics approval, Yuan Yang for sharing his knowledge of NARX models and Winfred Mugge for his critical notes during the entire process. Lastly, I'd like to thank the participants for partaking in this study.

## Appendices

### A. Model term section and parameter estimation

In this subsection, the procedures used for model term selection and parameter estimation are described.

#### *a. Model term selection*

The Forward Regression Orthogonal Least Squares (FROLS) algorithm [26] was used for model term selection, since it's computationally very efficient. Consider the full NARX structure as expressed in equation 3 reformulated in matrix form:

$$\boldsymbol{y} = P_0\boldsymbol{\Theta} + \boldsymbol{e} \tag{16}$$

with

$$P_0 = \begin{bmatrix}\boldsymbol{p_1} & \boldsymbol{p_2} & \dots & \boldsymbol{p_M}\end{bmatrix}$$

$$= \begin{bmatrix} p_1(1) & p_2(1) & \dots & p_M(1) \\ p_1(2) & p_2(2) & \dots & p_M(2) \\ \vdots & \vdots & & \vdots \\ p_1(N) & p_2(N) & \dots & p_M(N) \end{bmatrix} \tag{17}$$

where matrix $P_0$ contains all possible $M$ model terms (both first- and second-order) through time ($t = 1, 2, ..., N$); vector $\boldsymbol{\Theta}$ contains the model parameters; vector $\boldsymbol{y}$ represents the output signal, and vector $\boldsymbol{e}$ represents the model prediction error. The FROLS algorithm is an iterative method. In each step, one term $\boldsymbol{p_i}$ is added to the model, starting with the most prominent term.

Adding multiple model terms that explain the same portion of $\boldsymbol{y}$ does not improve the quality of the model, but increases its complexity. To avoid this scenario, potential model terms are judged based on the vector component orthogonal to the already selected terms, since this part of the vector can add new information to the model. The classical Gram-Schmidt algorithm was used to perform the orthogonalization [25].

The importance of each term was assessed based on the error-reduction-ratio (ERR). The ERR represents the fraction of the variance of the output vector, that can potentially be explained based on model term $p_i$.

$$ERR_i = \frac{\langle \boldsymbol{y}, \boldsymbol{w_i} \rangle^2}{\langle \boldsymbol{y}, \boldsymbol{y} \rangle \langle \boldsymbol{w_i}, \boldsymbol{w_i} \rangle} \tag{18}$$

where $\boldsymbol{y}$ represents the output vector; $\boldsymbol{w_i}$ represents a the orthogonal part of potential model term $\boldsymbol{p_i}$, and $\langle \boldsymbol{y}, \boldsymbol{w_i} \rangle$ denotes the inner product of vector $\boldsymbol{y}$ and vector $\boldsymbol{w_i}$.

The process of adding terms continues until the residual prediction error is smaller than a specified threshold.

*b. Parameter estimation*

In addition to selecting proper model terms, the FROLS algorithm performs a first parameter estimation (i.e. finds plausible values for $\theta$ in equation (3) for selected terms). Equation (16) can be rewritten after the Gram-Schmidt orthogonalization as [26]:

$$\boldsymbol{y} = P\boldsymbol{\Theta} + \boldsymbol{e} = W\boldsymbol{g} + \boldsymbol{e} \tag{19}$$

where matrix $P$ contains the selected original model terms and $\boldsymbol{\Theta}$ contains the corresponding parameters. Matrix $W$ in the second statement represents the orthogonalized components of the original terms, with their corresponding parameters in vector $\boldsymbol{g_i}$.

The advantage of orthogonalizing the model terms during each iteration is that their contributions are separated. Therefore, the parameter of each orthogonalized model term can be estimated separately based the projection of output vector $\boldsymbol{y}$ on orthogonalized model term $\boldsymbol{w_i}$:

$$\mathrm{proj}_{w_i \leftarrow y} = \underbrace{\frac{\langle \boldsymbol{y}, \boldsymbol{w_i} \rangle}{\langle \boldsymbol{w_i}, \boldsymbol{w_i} \rangle}}_{g_i} \boldsymbol{w_i} \tag{20}$$

where $g_i$ is the parameter for orthogonal model term $\boldsymbol{w_i}$. By reversing the classical Gram-Schmidt orthogonalization, original parameter vector $\boldsymbol{\Theta}$ can be retrieved from $\boldsymbol{g}$.

The FROLS algorithm results in a NARX model structure: the output is modelled based on previous input and output terms, but there are no past noise terms included in the model. Not including past noise terms before parameter estimation might bias the parameters. Therefore, the prediction error was computed based on the fitted NARX model and a linear noise model was added to matrix $P$ (equation (19)):

$$P = \begin{bmatrix} \boldsymbol{p_1} & \boldsymbol{p_2} & \dots & \boldsymbol{p_j} & \boldsymbol{e_{k-1}} & \boldsymbol{e_{k-2}} \end{bmatrix} \tag{21}$$

$$= \begin{bmatrix} p_1(1) & p_2(1) & \dots & p_j(1) & e_{k-1}(1) & e_{k-2}(1) \\ p_1(2) & p_2(2) & \dots & p_j(2) & e_{k-1}(2) & e_{k-2}(2) \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ p_1(N) & p_2(N) & \dots & p_j(N) & e_{k-1}(N) & e_{k-2}(N) \end{bmatrix}$$

where $p_1, ..., p_j$ are the selected model terms and vector $e_{k-1}$ and $e_{k-2}$ are the lagged prediction errors. Parameters were re-estimated based on equation (21) using a linear least squares approximation. After parameter estimation, the noise terms were discarded again.

**B. Recursive computation frequency mapping**

The polynomial NARX model is fitted to the data in the time-domain. However, neural oscillations can be more intuitively studied in the frequency domain. For this purpose, the parametric time domain model, in the shape of a difference equation, should be converted to the frequency domain. Peyton-Jones and Billings [27] analytically derived an algorithm to make this mapping possible. The algorithm and its derivation are described in detail in Billings' book [25]. To present a complete picture, the algorithm is summarized here in short.

To start, it is practical to highlight the difference between a 'model variable' and 'a model term'. A *model variable* is a previous instance of either the input signal $x_1$ or the output signal $x_2$. The *model terms* are build up from these model variables (e.g. $x_1(k-1)x_1(k-2)$ or $x_2(k-5)$). Note that $x_2(k-5)$ can be both a model variable and a model term, depending on how it's used.

The FROLS algorithm, described in Appendix A, resulted in a polynomial NARX model in the form of equation (3) and (4). These two equations can be combined and simplified. The time-domain NARX model can then be described as [25]:

$$x_1(k) = \tag{22}$$

$$\sum_{m=1}^{M} \sum_{p=0}^{m} \sum_{l_1,\dots,l_m=1}^{K} c_{pq}(l_1, ...., l_{p+q}) \prod_{i=1}^{p} x_1(k-l_i) \times \prod_{i=p+1}^{p+q} x_2(k-l_i)$$

In this representation, $M$ is the maximum order of nonlinearity, $K$ is the maximum time lag, $l_i$ is the time lag corresponding to model variable $i$, and structure $c_{pq}$ contains all model parameters (i.e. $\theta$ in equation (3)). Subscripts $p$ and $q$ denote the number of output and input model variables in the model term. For example, $c_{2,0}(2,4)$ corresponds to $\theta_{x_1(k-2)x_1(k-4)}$. If a model term is not part of the model, $c_{pq}$ is equal to zero.

The recursive algorithm by Peyton-Jones and Billings [27] was based on the form described in equation (22). Their goal was to derive an expression for $H_n$, where $n$ is the order of nonlinearity. The n-th order transfer function $H_n$ can be expressed as [21]:

$$H_n(f_1, ..., f_n) = \tag{23}$$

$$\frac{H_{n[x]}(f_1, ..., f_n) + H_{n[y]}(f_1, ..., f_n) + H_{n[xy]}(f_1, ..., f_n)}{1 - \sum_{l_1=1}^{K} c_{1,0}(l_1) e^{-j2\pi(f_1+...+f_n)l_1/f_s}}$$

where a distinction is made between contributions of pure input terms ($H_{n[x]}$), pure output terms ($H_{n[y]}$) and cross-terms ($H_{n[xy]}$), and $f_s$ is the sampling frequency. These separate contributions are defined as [21]:

$$H_{n[x]}(f_1, ..., f_n) =$$

$$\sum_{l_1, l_n=1}^{K} c_{0,n}(l_1, ..., l_n) e^{j2\pi(f_1 l_1+...+f_n l_n)/f_s} \tag{24}$$

$$H_{n[y]} =$$

$$\sum_{p=2}^{n} \sum_{l_1, l_n=1}^{K} c_{0,n}(l_1, ..., l_n) H_{n,p}(f_1, ..., f_n) \tag{25}$$

$$H_{n[xy]}(f_1, ..., f_n) =$$

$$\sum_{q=1}^{n-1} \sum_{p=1}^{n-q} \sum_{l_1, l_n=1}^{K} c_{p,q}(l_1, ..., l_n) \times H_{n-q,p}(f_1, ..., f_{n-q}) \times ...$$

$$... \, e^{-j2\pi(f_{n-q+1} l_{n-q+1}+...+f_{p+q} l_{p+q})/f_s}. \tag{26}$$

The algorithm is recursive since higher-order transfer functions depend on lower-order transfer functions. The recursion is defined as follows:

$$H_{n,p}(f_1, ..., f_n) = \tag{27}$$

$$\sum_{i=1}^{n-p+1} H_i(f_1, ..., f_i) H_{n-i,p-1}(f_{i+1}, ..., f_n) e^{-j2\pi(f_1+...+f_i)l_p/f_s},$$

and, eventually,

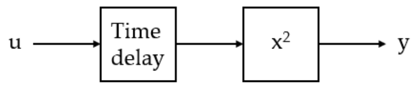
$$H_{n,1}(f_1, ..., f_n) = H_n(f_1, ..., f_n) e^{-j2\pi(f1+...+f_n)l_1 f_s}. \tag{28}$$

We implemented this algorithm in MatLab 2015b, since there was no existing package available.

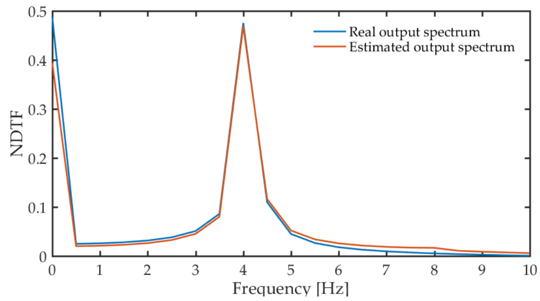
**C. Effect of discrete approximation on predicted output spectrum**

This appendix contains a visual example to understand the effect of the discrete approximation on the predicted output spectrum.

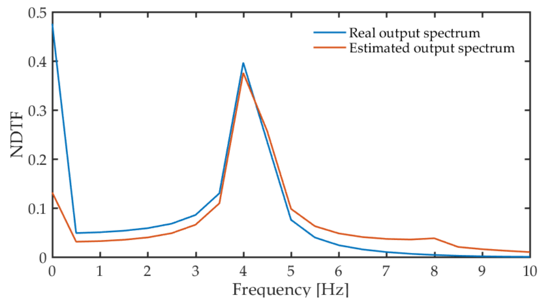
For this example, the simple open-loop system $y(k) = u(k-1)u(k-1)$ was used. The system was run for 2 seconds at a 20 Hz sampling frequency. Since it is an open-loop system without noise, the NDTF should approach the real output spectrum $|Y|$.




If a 2 Hz sinus is used as input $u$, which is part of the frequency vector, the NDTF mimics the real output spectrum perfectly.

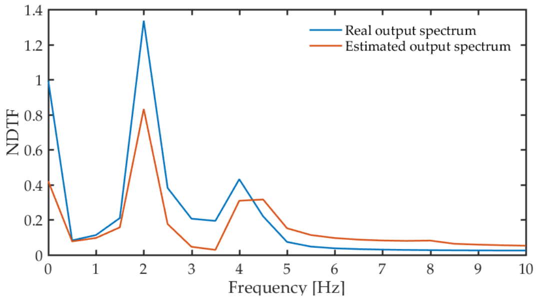


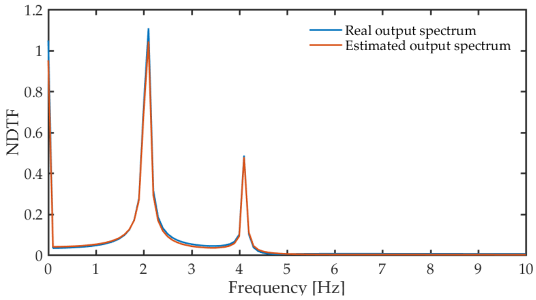

However, when the frequency of the input vector is shifted to 2.05 Hz, which is not part of the frequency vector, the NDTF starts to deviate.



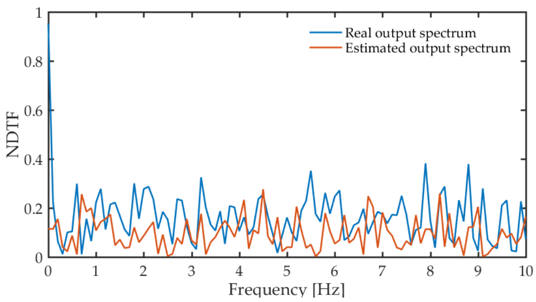

The deviating behavior becomes worse if a multisine is used as the input, with frequency content on 2.05 Hz and 3.13 Hz.

The discrete approximation, logically, becomes better for a higher frequency resolution. If the recording time is increased from 2 seconds to 10 seconds, the NDTF performs well again, even though 2.05 and 3.13 are still not in the frequency vector.




Lastly, a white noise signal was used as the input, while keeping the high frequency resolution.




In the last figure, the effect of the discretization starts to become clear. The orange line clearly lies below the the blue line. The NDTF underestimates the actual causal influence, since not all frequency combinations are taken into account.

## References


[1] X.-J. Wang, "Neurophysiological and computational principles of cortical rhythms in cognition," *Physiological reviews*, vol. 90, no. 3, pp. 1195–1268, 2010.

[2] B. van Wijk, P. J. Beek, and A. Daffertshofer, "Neural synchrony within the motor system: what have we learned so far?," *Frontiers in human neuroscience*, vol. 6, p. 252, 2012.

[3] B. Conway, D. Halliday, S. Farmer, U. Shahani, P. Maas, A. Weir, and J. Rosenberg, "Synchronization between motor cortex and spinal motoneuronal pool during the performance of a maintained motor task in man.," *The Journal of physiology*, vol. 489, no. 3, pp. 917–924, 1995.

[4] J. Ushiyama, T. Suzuki, Y. Masakado, K. Hase, A. Kimura, M. Liu, and J. Ushiba, "Between-subject variance in the magnitude of corticomuscular coherence during tonic isometric contraction of the tibialis anterior muscle in healthy young adults," *Journal of neurophysiology*, vol. 106, no. 3, pp. 1379–1388, 2011.

[5] J. Groß, P. Tass, S. Salenius, R. Hari, H.-J. Freund, and A. Schnitzler, "Cortico-muscular synchronization during isometric muscle contraction in humans as revealed by magnetoencephalography," *The Journal of Physiology*, vol. 527, no. 3, pp. 623–631, 2000.

[6] Y. Fang, J. J. Daly, J. Sun, K. Hvorat, E. Fredrickson, S. Pundik, V. Sahgal, and G. H. Yue, "Functional corticomuscular connection during reaching is weakened following stroke," *Clinical neurophysiology*, vol. 120, no. 5, pp. 994–1002, 2009.

[7] J. N. Caviness, H. A. Shill, M. N. Sabbagh, V. G. Evidente, J. L. Hernandez, and C. H. Adler, "Corticomuscular coherence is increased in the small postural tremor of parkinson's disease," *Movement disorders*, vol. 21, no. 4, pp. 492–499, 2006.

[8] S. Salenius, S. Avikainen, S. Kaakkola, R. Hari, and P. Brown, "Defective cortical drive to muscle in parkinsons disease and its improvement with levodopa," *Brain*, vol. 125, no. 3, pp. 491–500, 2002.

[9] S. F. Campfens, H. van der Kooij, and A. C. Schouten, "Face to phase: pitfalls in time delay estimation from coherency phase," *Journal of computational neuroscience*, vol. 37, no. 1, pp. 1–8, 2014.

[10] C. N. Riddle and S. N. Baker, "Manipulation of peripheral neural feedback loops alters human corticomuscular coherence," *The Journal of physiology*, vol. 566, no. 2, pp. 625–639, 2005.

[11] M. P. Vlaar, T. Solis-Escalante, A. N. Vardy, F. C. Van Der Helm, and A. C. Schouten, "Quantifying nonlinear contributions to cortical responses evoked by continuous wrist manipulation," *IEEE transactions on neural systems and*

*rehabilitation engineering*, vol. 25, no. 5, pp. 481–491, 2017.

[12] M. P. Mileusnic, I. E. Brown, N. Lan, and G. E. Loeb, "Mathematical models of proprioceptors. i. control and transduction in the muscle spindle," *Journal of neurophysiology*, vol. 96, no. 4, pp. 1772–1788, 2006.

[13] R. Pintelon and J. Schoukens, *System identification: a frequency domain approach*. John Wiley & Sons, 2012.

[14] Y. Yang, T. Solis-Escalante, J. Yao, A. Daffertshofer, A. C. Schouten, and F. C. Van Der Helm, "A general approach for quantifying nonlinear connectivity in the nervous system based on phase coupling," *International journal of neural systems*, vol. 26, no. 01, p. 1550031, 2016.

[15] Y. Yang, A. C. Schouten, T. Solis-Escalante, and F. C. van der Helm, "Probing the nonlinearity in neural systems using cross-frequency coherence framework," *IFAC-papersonline*, vol. 48, no. 28, pp. 1386–1390, 2015.

[16] S. F. Campfens, A. C. Schouten, M. J. van Putten, and H. van der Kooij, "Quantifying connectivity via efferent and afferent pathways in motor control using coherence measures and joint position perturbations," *Experimental brain research*, vol. 228, no. 2, pp. 141–153, 2013.

[17] K. J. Blinowska, "Review of the methods of determination of directed connectivity from multichannel data," *Medical & biological engineering & computing*, vol. 49, no. 5, pp. 521–529, 2011.

[18] M. Kaminski and K. J. Blinowska, "A new method of the description of the information flow in the brain structures," *Biological cybernetics*, vol. 65, no. 3, pp. 203–210, 1991.

[19] L. A. Baccalá and K. Sameshima, "Partial directed coherence: a new concept in neural structure determination," *Biological cybernetics*, vol. 84, no. 6, pp. 463–474, 2001.

[20] S. Chen and S. A. Billings, "Representations of non-linear systems: the narmax model," *International Journal of Control*, vol. 49, no. 3, pp. 1013–1032, 1989.

[21] F. He, S. A. Billings, H.-L. Wei, and P. G. Sarrigiannis, "A nonlinear causality measure in the frequency domain: Nonlinear partial directed coherence with applications to eeg," *Journal of neuroscience methods*, vol. 225, pp. 71–80, 2014.

[22] E. Ribot-Ciscar, V. Hospod, J.-P. Roll, and J.-M. Aimonetti, "Fusimotor drive may adjust muscle spindle feedback to task requirements in humans," *Journal of neurophysiology*, vol. 101, no. 2, pp. 633–640, 2009.

[23] W. Mugge, J. Schuurmans, A. C. Schouten, and F. C. van der Helm, "Sensory weighting of force and position feedback in human motor control tasks," *Journal of Neuroscience*, vol. 29, no. 17, pp. 5476–5482, 2009.

[24] C. W. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica: Journal of the Econometric Society*, pp. 424–438, 1969.

[25] S. A. Billings, *Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains*. John Wiley & Sons, 2013.

[26] S. Chen, S. A. Billings, and W. Luo, "Orthogonal least squares methods and their application to non-linear system identification," *International Journal of control*, vol. 50, no. 5, pp. 1873–1896, 1989.

[27] J. P. Jones and S. Billings, "Recursive algorithm for computing the frequency response of a class of non-linear difference equation models," *International Journal of Control*, vol. 50, no. 5, pp. 1925–1940, 1989.

[28] S. Potgieter, "Sensitivity analysis of (renormalized) partial directed coherence for the study of corticomuscular causality," 2013.

[29] T. Mima, T. Matsuoka, and M. Hallett, "Information flow from the sensorimotor cortex to muscle in humans," *Clinical Neurophysiology*, vol. 112, no. 1, pp. 122–126, 2001.

[30] C. L. Witham, C. N. Riddle, M. R. Baker, and S. N. Baker, "Contributions of descending and ascending pathways to corticomuscular coherence in humans," *The Journal of physiology*, vol. 589, no. 15, pp. 3789–3800, 2011.

[31] K. Krishnanathan, S. R. Anderson, S. A. Billings, and V. Kadirkamanathan, "A data-driven framework for identifying nonlinear dynamic models of genetic parts," *ACS synthetic biology*, vol. 1, no. 8, pp. 375–384, 2012.

[32] M. A. Balikhin, R. J. Boynton, S. N. Walker, J. Borovsky, S. A. Billings, and H. Wei, "Using the narmax approach to model the evolution

of energetic electrons fluxes at geostationary orbit," *Geophysical Research Letters*, vol. 38, no. 18, 2011.

[33] E. R. Williams and S. N. Baker, "Circuits generating corticomuscular coherence investigated using a biophysically based computational model. i. descending systems," *Journal of neurophysiology*, vol. 101, no. 1, pp. 31–41, 2009.

[34] S. Heitmann, T. Boonstra, P. Gong, M. Breakspear, and B. Ermentrout, "The rhythms of steady posture: Motor commands as spatially organized oscillation patterns," *Neurocomputing*, vol. 170, pp. 3–14, 2015.

[35] S. F. Campfens, S. B. Zandvliet, C. G. Meskers, A. C. Schouten, M. J. van Putten, and H. van der Kooij, "Poor motor function is associated with reduced sensory processing after stroke," *Experimental brain research*, vol. 233, no. 4, pp. 1339–1349, 2015.

[36] L. Timmermann, J. Gross, M. Dirks, J. Volkmann, H.-J. Freund, and A. Schnitzler, "The cerebral oscillatory network of parkinsonian resting tremor," *Brain*, vol. 126, no. 1, pp. 199–212, 2002.

[37] F. J. Harris, "On the use of windows for harmonic analysis with the discrete fourier transform," *Proceedings of the IEEE*, vol. 66, no. 1, pp. 51–83, 1978.

[38] Y. Yang, T. Solis-Escalante, M. van de Ruit, F. C. van der Helm, and A. C. Schouten, "Nonlinear coupling between cortical oscillations and muscle activity during isotonic wrist flexion," *Frontiers in computational neuroscience*, vol. 10, p. 126, 2016.

[39] M. Lim, J. S. Kim, M. Kim, and C. K. Chung, "Ascending beta oscillation from finger muscle to sensorimotor cortex contributes to enhanced steady-state isometric contraction in humans," *Clinical Neurophysiology*, vol. 125, no. 10, pp. 2036–2045, 2014.

[40] F. Meng, K.-y. Tong, S.-t. Chan, W.-w. Wong, K.-h. Lui, K.-w. Tang, X. Gao, and S. Gao, "Study on connectivity between coherent central rhythm and electromyographic activities," *Journal of neural engineering*, vol. 5, no. 3, p. 324, 2008.

[41] H.-L. Wei and S. A. Billings, "Model structure selection using an integrated forward orthogonal search algorithm assisted by squared correlation and mutual information," *International Journal of Modelling, Identification and Control*, vol. 3, no. 4, pp. 341–356, 2008.

[42] J. Peyton Jones and S. Billings, "Interpretation of non-linear frequency response functions," *International Journal of Control*, vol. 52, no. 2, pp. 319–346, 1990.

[43] S. Chen, E. Chng, and K. Alkadhimi, "Regularized orthogonal least squares algorithm for constructing radial basis function networks," *International Journal of Control*, vol. 64, no. 5, pp. 829–837, 1996.

[44] S. BILLINGS and W. Voon, "Correlation based model validity tests for non-linear models," *International journal of Control*, vol. 44, no. 1, pp. 235–244, 1986.

[45] E. Florin, J. Gross, J. Pfeifer, G. R. Fink, and L. Timmermann, "Reliability of multivariate causality measures for neural data," *Journal of neuroscience methods*, vol. 198, no. 2, pp. 344–358, 2011.

[46] T.-P. Jung, S. Makeig, C. Humphries, T.-W. Lee, M. J. Mckeown, V. Iragui, and T. J. Sejnowski, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology*, vol. 37, no. 2, pp. 163–178, 2000.

TU Delft