Synchronizing Decentralized Control Loops for Overall Performance Enhancement: A Youla Framework Applied to a Wafer Scanner

Enzo Evers * Marc van de Wal ** Tom Oomen *

* Eindhoven University of Technology, Eindhoven, The Netherlands. (e-mail: E.Evers@tue.nl)
** ASML Research, Veldhoven, The Netherlands.

Abstract: Manufacturing equipment often consists of multiple subsystems. For instance, in lithographic IC manufacturing, both a reticle stage and a wafer stage move synchronously. Traditionally, these subsystems are divided into manageable subproblems. This leads to suboptimal performance. In particular, limited by the performance limitations (Seron et al. 1997) of all subsystems. An example of a manufacturing machine in the high-tech industry with multiple positioning subsystems is a wafer scanner.

In wafer scanners multiple positioning systems move simultaneously to produce ICs. In particular, to expose each die, the reticle stage (RS) and wafer stage (WS) move opposite to each other in a scanning motion. Accuracy during exposure depends on the relative error between RS and WS. Traditionally, both stages track the same, appropriately scaled for optical magnification, trajectory. Therefore, minimization of the two individual tracking errors is a sufficient condition for a small relative error. However, it is not necessary and significantly limited by the performance limitations of the individual subsystem. As an example, the subsystems generally contain NMP zeros. In contrast, the overall, essentially non-square system, generally have no NMP zeros at all, as is shown in this paper.

Several attempts have been made to address overall system performance. Research examples of improved synchronization by advanced feed-forward can be found in Navarrete et al. (2015), see also Boeren et al. (2014) for related feed-forward results. In Barton and Alleyne (2008) and Mishra et al. (2008) improved synchronization is achieved by using iterative learning control. Rational filters are used in direct feedback of the relative error in Wang et al. (2006). Typically, in these experimental studies the controller is extended by an uni-directional interaction, hereafter referred to as uni-directional Feed-Through (FT). Such uni-directional interaction is also considered in Sakata and Fujimoto (2009) using rational filters and Heertjes and Tenmizer (2012) using data-based optimized FIR filters. Uni-directional coupling from the WS to the RS enables a significant performance improvement. A secondary coupling from the RS to the WS allows for inherently better performance due to a larger design freedom. Yet it is often not applied, as bi-directional coupling affects closed-loop stability (Skogestad and Postlethwaite 2009), which is not the case for uni-directional coupling or interference.

Although several attempts to improve overall system performances have been made, present solutions are ad hoc. The aim of this paper is to design a systematic approach, consisting of 1) parameterization of all admissible solutions, 2) a stability proof and 3) design guidelines for the coupling elements.

Alternatively, improved overall system performance can be achieved by designing a full MIMO (Multi-In Multi-Out) replacement controller. However, the system models needed for such an approach are complex and not trivially obtained. Therefore, the current perturbed controller approach allows for a control relevant add-on element to the existing decentralized control architecture. The approach

Keywords: Mechatronic systems, Motion Control Systems, Decentralized control, Coupling

Machines used in the manufacturing industry often consist of multiple subsystems. In view of complexity, the design and control of these machines is divided into manageable subproblems. This leads to suboptimal performance. In particular, limited by the performance limitations (Seron et al. 1997) of all subsystems. An example of a manufacturing machine in the high-tech industry with multiple positioning subsystems is a wafer scanner.
The main contributions of this paper are threefold:

1. A framework for achieving overall machine performance through bi-directional coupling.
2. Design guidelines for controller tuning, both norm-optimal ($H_2$, $H_{\infty}$) and manual tuning.
3. Case study using measurement data of an industrial wafer scanner.

To facilitate presentation the framework is developed for a $2 \times 2$ scalar case, and notation is used for a wafer scanner to avoid confusion. The framework is applicable to the general MIMO situation for a large number of applications.

1. CONTROL PROBLEM ANALYSIS & REQUIREMENTS FOR BI-DIRECTIONAL COUPLING.

The aim of this paper is to construct a framework for bi-directional controller coupling, using a case-study based on an industrial wafer scanner. The following section describes the setup and motivation for this approach followed by the problem definition and requirements for bi-directional coupling.

1.1 Setup

The uncoupled control diagram for the reticle and wafer stage is shown in Fig. 1. In particular, except for the reference signal, the loops are not connected. The subsystems $P_r$ and $P_w$ are externally approximately decoupled in the full system due to an excellent mechanical design. Both individual subsystems typically contain 6 Degrees Of Freedom (DOF) resulting in a $(6 \times 6)$ system model. Typically, both subsystems can be considered approximately decoupled (Oomen and Steinbuch 2017). Thus, in the scan direction, the combined system can be modeled as a $(2 \times 2)$ diagonal plant and controller, i.e.

$$ P_0 = \begin{bmatrix} P_r & 0 \\ 0 & P_w \end{bmatrix}, \quad K_0 = \begin{bmatrix} C_r & 0 \\ 0 & C_w \end{bmatrix} \quad (1) $$

where $P_0$ and $K_0$ are the nominal plant and controller parameterizations respectively. In this paper, the main goal is to design an add-on control element that retains the existing diagonal control elements and minimizes $e_{wr} = y_w - \gamma^{-1} y_r$. Throughout, it is often tacitly assumed that the sub-blocks in (1) are scalar to facilitate the presentation. Many results directly apply or can be extended to the MIMO case.

1.2 Motivation: beyond traditional performance limitations.

The additional controller freedom can be used to achieve combined system performance beyond the limitations of the decentralized ($2 \times $ Single-In Single-Out (SISO)) individual subsystems. One of these limitations is nonminimum-phase (NMP) zeros. These zeros can be caused by non-collocated sensor and actuator placement (Hong and Bernstein 1998), as well as sampling (Åström et al. 1984) of continuous time systems. Their limiting effect on the achievable performance can be seen by using a Poisson integral relation (Freudenberg et al. 2003). Consider the system in Fig. 1 that is represented using the standard plant notation, as shown in Fig. 2, as a disturbance attenuation problem. Where $y = [e_r, e_w]^T$, $u = [v_r, v_w]^T$ and all disturbances are represented as two unique output disturbances $w = [v_r, v_w]^T$ such that $v_r$ and $v_w$ are uncorrelated. The performance variable $z$ is chosen as either 1) traditional $z_1 = [e_r, e_w]^T$ or 2) as considered here, $z_2 = e_{wr} = [\gamma^{-1} e_r - e_w]$.  

1. Using $z_1$

$$ \begin{bmatrix} e_r \\ e_w \end{bmatrix} = \begin{bmatrix} S_r \\ S_w \end{bmatrix} \begin{bmatrix} S_w \\ S_w \end{bmatrix} \begin{bmatrix} v_r \\ v_w \end{bmatrix} T_{zw_1} \quad (2) $$

2. Using $z_2$

$$ e_{wr} = -[\gamma^{-1} - I] \begin{bmatrix} S_r \\ S_w \end{bmatrix} \begin{bmatrix} S_r \\ S_w \end{bmatrix} \begin{bmatrix} v_r \\ v_w \end{bmatrix} T_{zw_2} $$

where $S_r = (I + P_r C_r)^{-1}$ and $S_w = (I + P_w C_w)^{-1}$ are fixed and $S_{rw}$ and $S_{wr}$ represent possible coupling introduced by FT. The two individual stages are constructed and optimized as stand-alone systems. This follows from a long standing industrial practice of decentralized control. Therefore, the controllers $C_r$ and $C_w$ are considered fixed and additional controller freedom can only be added by inter-connecting the existing subsystems. Here, both $S_r$ and $S_w$ are optimized to achieve the desired disturbance attenuation w.r.t. $v_r$ and $v_w$.

Using $z_1$ yields the system $T_{zw_1}$ and any other choice than $S_{rw} = S_{wr} = 0$ is clearly undesired as it introduces additional disturbance in $e_r$ and $e_w$. This shows that: 1) the additional controller freedom that is added by FT, in the form of $S_{rw}$ and $S_{wr}$, cannot improve traditional
performance beyond the individual limits imposed by the subsystems. And 2) that any additional coupling introduced by FT only deteriorates traditional performance. However, using z as the system $T_{zw}$ is potentially no longer limited by the subsystems as $S_w$ and $S_{zw}$ can be used to achieve disturbance suppression in regions where $S$ and $S_w$ cannot. This retains the original control characteristics while still allowing for improved performance.

Achieving performance beyond individual stage limits can be transparently done by connecting the additional add-on controller freedom with the new performance criterion, which is proposed in this paper.

1.3 Problem definition

Taking into account the various constraints and desired system characteristics the FT design problem is formulated as follows:

**Definition 1.** (FT design problem). Minimize $e_{aw}$ while ensuring that:

(i) The original closed-loop transfer functions of each individual subsystem remains unchanged.
(ii) The controller input signals, the signals entering $C_r$ and $C_w$ in Fig. 1, are invariant to FT.
(iii) Well-posedness is guaranteed.
(iv) Nominal internal stability is guaranteed.
(v) The FT filters are real rational and stable.

The interpretation of the above is as follows. Req. (i) ensures an add-on design, i.e. the added FT must be complementary to the existing control structure by ensuring that the original decentralized closed-loop characteristics are preserved. Req. (ii) is closely related to (i). It originates from practical considerations to avoid required re-tuning of the RS and WS loop. Here, invariance is achieved if the controller input signals are not changed by the additional controller freedom created by the additional FT filters. Req. (iii) and (iv) are necessary to retain a well-posed and internally stable feedback system. Req. (v) offers robustness against re-tuning and loop failures and thereby facilitates industrial implementation. A system $G$ that is real rational and stable, i.e. has no unstable poles, will be denoted as $G \in RH_{\infty}$. Restricting the FT filters to $RH_{\infty}$ will turn out to significantly simplify the filter design.

2. A YOULA APPROACH TO BI-DIRECTIONAL FT

Bi-directional interaction is inherently complex and properties such as well-posedness and internal stability necessitate a detailed analysis (Maciejowski 1989). Therefore, the bi-directional FT problem is recast as a Youla parameterization using coprime factorization, shown in Fig. 3. The Youla parameterization allows for a straightforward analysis and separation of the nominal controller $K_0 = N_k D_k^{-1}$ and the additional add-on controller freedom.

The Youla parameterization (Youla et al. 1976) describes the set of all possible controllers that stabilize a nominal plant $P_0 \in \mathcal{R}$ as a function of a nominal controller $K_0 \in \mathcal{R}$ and a Youla parameter $\Delta_k \in RH_{\infty}$. Both the nominal plant $P_0 = ND^{-1}$ and nominal controller $K_0 = N_k D_k^{-1}$ are represented by right-coprime factorizations.

Fig. 3. Controller diagram using a coprime Youla parameterization. Here, all systems are $(2 \times 2)$ transfer function matrices for each DOF.

A similar formulation can be formed using left-coprime factorization.

**Definition 2.** (Right-coprime factorization (rcf)). The ordered pair $\{N, D\}$, with $D \in RH_{\infty}^q$ and $N \in RH_{\infty}^p$, is a right-coprime factorization (rcf) of $P \in RH_{\infty}$ if

\[
\begin{align*}
&i \quad D \text{ is invertible (square and non-singular),} \\
&ii \quad P = ND^{-1}, \\
&iii \quad N \text{ and } D \text{ are right-coprime.}
\end{align*}
\]

$N$ and $D$ are right-coprime if there exist matrices $W_r, L_r \in RH_{\infty}$ such that the Bezout identity (Zhou et al. 1996)

\[
L_r N + W_r D = I,
\]

holds

**Theorem 3.** (Set of stabilizing controllers).

Let $P_0 = ND^{-1}$ and $K_0 = N_k D_k^{-1}$ where $\{N, D\}$, $\{N_k, D_k\}$ are rcf of $P_0$, $K_0$. Let perturbed controller factors be defined as

\[
N_{k\Delta} := N_k + D\Delta_k, \quad D_{k\Delta} := D_k - N\Delta_k,
\]

such that

\[
K_{\Delta} = N_{k\Delta} D_{k\Delta}^{-1}
\]

where $\Delta_k$ are the free Youla parameters. Then it follows that $K_{\Delta}$ stabilizes $P_0$ iff $\Delta_k \in RH_{\infty}$.

For a proof see, e.g., (Schrama et al. 1992).

Stability follows because the closed-loop system $F_l(P_0, K_{\Delta})$ can be written as the sum of the original closed-loop system $F_l(P_0, K_0)$, which is assumed to be stable, and an additional factor that is affine in $\Delta_k \in RH_{\infty}$. The main idea is to cast the bi-directional FT problem into Youla form by using a suitable coprime factorization and the result in Thm. 3.

Assuming $K_0 \in RH_{\infty}$, then

\[
K_0 = \begin{bmatrix} C_r & 0 \\ 0 & C_w \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}^{-1},
\]

where $N_k, D_k \in RH_{\infty}$, is a suitable RCF. If $C_r, C_w$ have integral control action, i.e., poles on the imaginary axis, care has to be taken. Pragmatically, these controller poles can be shifted slightly into the left half of the complex plane to satisfy $K_0 \in RH_{\infty}$. Similarly, the nominal plant $P_0$ can be written as

\[
P_0 = \begin{bmatrix} Z_r & 0 \\ 0 & Z_w \end{bmatrix} \begin{bmatrix} P_{r}^{-1} Z_r & 0 \\ 0 & P_{w}^{-1} Z_w \end{bmatrix}^{-1},
\]
such that \( N, D \in \mathcal{RH}_\infty \). For (9) to be a coprime factorization, it is required that both \( P_r, P_w \) are minimal realizations and \( Z_r, Z_w \) do not contain additional RHP zeros other than those required to ensure \( D \in \mathcal{RH}_\infty \). Then, define \( \Delta_k \) as
\[
\Delta_k = \begin{bmatrix} 0 & \hat{X} \\hat{Y} & 0 \end{bmatrix} \in \mathcal{RH}_\infty,
\]
where \( \hat{X} \) and \( \hat{Y} \) are both free-design parameters. The set of stabilizing controllers (7) then becomes
\[
K_{\Delta_k} = \begin{bmatrix} C_r \quad P_r^{-1} Z_r \hat{X} \quad I \quad -Z_r \hat{X} \quad -Z_w \hat{Y} \quad I \end{bmatrix}^{-1} \cdot \quad (11)
\]
Using this stabilizing controller the desired system properties are achieved, described by the following theorem.

**Theorem 4.** (Nominal bi-directional FT). Given a feed-back loop where the nominal controller \( K_0 \) and plant \( P_0 \) are described by (8) and (9), respectively, and \( \Delta_k \) is chosen as in (10), then the set of stabilizing bi-directional FT controllers for \( P_0 \) is given by:
\[
K_{\Delta_k}(P_0) = \{ N_k D_k^{-1} | \Delta_k \in \mathcal{RH}_\infty \} \quad (13)
\]
Any controller \( K_{FT} = K_{\Delta_k}(P_0) \) in (13) then achieves:
(a) \( \forall \Delta_k, f_i(P_0, K_{FT}) \) is affine in \( \Delta_k \).
(b) \( \forall \Delta_k, f_i(P_0, K_{FT})_{ii} = f_i(P_0, K_0)_{ii} \), where \( i = [1,2] \) of each 2 \( \times \) 2 block matrix entry.
(c) \( \forall \Delta_k, e_i^r \neq f(e_i^w) \) and \( e_i^w \neq f(e_i^r) \), i.e. \( e_i^r, e_i^w \) are invariant to FT.
(d) \( \forall \Delta_k \in \mathcal{RH}_\infty, f_i(P_0, K_{FT}) \) is well-posed.
(e) \( \forall \Delta_k \in \mathcal{RH}_\infty, f_i(P_0, K_{FT}) \) is internally stable.

**Proof.** The proof follows by straightforward evaluation of the bi-directional closed-loop system \( f_i(P_0, K_{FT}) \) as shown in (12), where \( X = Z_r \hat{X} \) and \( Y = Z_w \hat{Y} \). The parameters \( X, Y \) are clearly affine in \( \hat{X}, \hat{Y} \) and \( Z_r, Z_w \in \mathcal{RH}_\infty \) by coprime factorization. Therefore, assuming that \( f_i(P_0, K_0) \) is internally stable and well-posed, any controller in (13) achieves an internally stable and well-posed closed-loop \( f_i(P_0, K_{FT}) \). Furthermore, the Youla parameterization in (13) clearly achieves (v) in Def. 1.

By using the stabilizing controller \( K_{\Delta_k}(P_0) \) the full transfer function matrix between the disturbances \( w \) and the performance variable \( z_2 = e_{wr} \) becomes
\[
e_{wr} = -(I + \gamma Y) [S_r S_P T_r][v_r \quad u_r] + [-\gamma Y S_r + T_r][\eta_r] + (I + \gamma^{-1} X) [S_w P S_w T_w][v_w \quad u_w] + [\gamma^{-1} X S_w - T_w][\eta_w]. \quad (14)
\]
It shows that the change in the closed-loop relations of the system disturbances to \( e_{wr} \) can be expressed by two factors: 1) \((I + \gamma^{-1} X)\) and 2) \(- (I + \gamma Y)\) that are both affine in \( X = Z_r \hat{X} \) and \( Y = Z_w \hat{Y} \). These factors will henceforth be referred to as the improvement factors. This will allow for straightforward design and analysis of the coupled closed-loop system.

### 3. SYSTEMATIC DESIGN APPROACH

It is shown in (14) that the change in closed-loop behavior induced by the additional controller coupling can be described by evaluating the improvement factors. In this section the limitations and design guidelines of the coupling filters are presented.

#### 3.1 Limitations

By evaluating the factors \((\gamma^{-1} X + I)\) and \((I + \gamma Y)\) in (14), the region of performance improvement or deterioration, can directly be determined. Assuming that the force actuator is capable of supplying infinite force to the system, the overall performance is limited by the inverse system model. When minimizing the improvement factor, the design parameters \( X \) and \( Y \) are ideally chosen as \( X = -\gamma I \) and \( Y = -\gamma^{-1} \). However, this generally leads to inadmissible FT filters \( \{H_{w2r} = P_r^{-1} X, H_{r2w} = P_w^{-1} Y\} \notin \mathcal{RH}_\infty \).

Typical causes include:

i) Non-Minimum Phase (NMP) zeros in the plant \( P_r \), since inversion would make \( P_r^{-1} \) unstable and thus \( P_r^{-1} \notin \mathcal{RH}_\infty \).

ii) I/O delay in the plant \( P_r \), since inversion would result in an a-causal filter, which cannot be implemented in real-time.

iii) Pole/zero excess in the plant \( P_r \), since inversion would make \( P_r^{-1} \) non-proper and thus \( P_r^{-1} \notin \mathcal{RH}_\infty \).

#### 3.2 Design guidelines

In the proposed framework the parameter \( X \) consists of two components, i.e., \( X := Z_r \hat{X} \): 1) a component denoted as \( Z_r \), that, in view of requirements, must be chosen such that \( P_r^{-1} X \in \mathcal{RH}_\infty \); 2) a component \( \hat{X} \) that can be designed to minimize \( e_{wr} \), e.g., \( |(\gamma^{-1} Z_r \hat{X} + I)| < 1 \) for specific, e.g., low, frequency ranges. Similar approaches for tracking control includes, e.g., (Tomizuka 1987). The following guidelines aim to design \( X := Z_r \hat{X} \) such that:
1) (9) is a valid RCF and 2) \( e_{wr} \) is minimized under the constraints posed in Def 1.

**Remark 5.** In this section \( P_r \) is used as an example, the design guidelines for \( P_w \) are comparable.

**Satisfying constraints** The NMP elements (i) and (ii) of \( P_r \) are duplicated in \( Z_r \) such that they cancel in \( P_r^{-1} Z_r \). However, this generally makes \( Z_r \hat{X} \notin \mathcal{RH}_\infty \). Therefore, the NMP elements in \( Z_r \) are constructed as all-pass, i.e., \( [Z_r, (j\omega)] = 1 \) \( \forall \omega \), to ensure that the parameter \( Z_r \hat{X} \in \mathcal{RH}_\infty \) as required by (9). The filter \( P_r^{-1} Z_r \) is made proper by including a low-pass filter with a high-frequency cutoff in \( Z_r \) of order equal to or greater than the amount of pole/zero excess (iii).

**Free design parameter** Using the free design parameter \( \hat{X} \) the improvement factor \((\gamma^{-1} Z_r \hat{X} + I)\) can be shaped such that the coupled closed-loop disturbance attenuation is improvement at the desired frequency region. The parameter can be shaped using manual loop-shaping or norm optimal techniques such as \( \mathcal{H}_2, \mathcal{H}_\infty \) controller design. If
\[
\begin{bmatrix}
y_r \\
y_w \\
e_r \\
e_w \\
u_r \\
u_w
\end{bmatrix} = \begin{bmatrix}
S_r & -XS_w & PS_r & -XPS_w & T_r & -XT_w & -T_r & -XS_w \\
-YS_r & S_u & -YS_p & PS_r & YT_r & XT_w & T_r & -TS_w \\
S_r & XS_w & -PS_r & XPS_w & YS_r & -YS_w & -TS_r & S_u \\
YS_r & -S_u & YF_s & -PS_u & -YS_r & YT_w & YS_r & -TS_u \\
0 & -S_u & 0 & -PS_u & S_r & 0 & 0 & -S_u \\
P^{-1}YS_r & CS_r & P^{-1}YS_pr & S_w & P^{-1}YTS_r & CS_r & P^{-1}XTS_w & CS_w
\end{bmatrix}
\begin{bmatrix}
v_r \\
v_w \\
d_p \\
d_w \\
\psi_r \\
\psi_w
\end{bmatrix}
\]

(12)

The improvement factors for R2W and W2R coupling. It shows that for low frequencies a significant improvement in disturbance attenuation can be achieved.

![Amplitude vs Frequency](image1.png)

![Amplitude vs Frequency](image2.png)

Fig. 4. The improvement factors for R2W and W2R coupling. It shows that for low frequencies a significant improvement in disturbance attenuation can be achieved.

the free design parameter is chosen as $\hat{X} = -\gamma I$, it is ensured that $|\gamma^{-1}X + I| << 1$ at low frequencies as long as $|Z_r| \approx 1$ and has the correct phase.

4. INDUSTRIAL CASE STUDY

The proposed method is applied to an industrial wafer scanner in a case study using nominal system models. Decentralized control is achieved by tightly tuned feedback controllers that are optimized in terms of control bandwidth for each of the subsystems.

By applying the design procedure for nominal bi-directional FT a set of stabilizing coupling filters is constructed. The result in Fig. 4 shows the improvement factors $(I + \gamma Y)$ (R2W) and $(I + \gamma^{-1}X)$ (W2R). The large dips at high frequencies are caused by the phase shift in the all-pass elements of $Z_r$ and $Z_w$, e.g., $\{\gamma Y|\gamma^{-1}X\} \in [-1, 1]$, causing $(I + \gamma Y)$ and $(I + \gamma^{-1}X)$ to locally approach 0. The coupled closed-loop transfer functions are shown in Fig. 5 as a direct affine combination of the original closed-loop transfer functions and the improvement factors shown in Fig. 4. It shows that using the additional controller freedom, superior disturbance suppression is achieved in the low-frequency region.

The time domain performance is investigated by analyzing 8 recorded servo error traces of a single die exposure. The expected performance using bi-directional FT is then simulated using the system models. The results in in Fig. 6 show that at low-frequencies, the RS is used to track the performance limited WS. It shows that low-frequency improvement in the disturbance attenuation comes at the cost of slight high-frequency deterioration. However, since the setpoint and disturbances contain most energy at low frequencies, this trade-off result in superior overall system performance. The results in Fig. 7 show that a significant improvement in the performance variable $e_{wr}$ is achieved at low-frequencies by approximately synchronizing the two stages.

5. CONCLUSION

A systematic approach to bi-directional controller coupling is proposed using a Youla parameterization with coprime factorization. Emphasis is placed on industrial applicability by designing the additional coupling as add-on, thus preserving the original closed-loop characteristics. The coupled system behavior is expressed as a straightforward combination of the original closed-loop characteristics and an affine factor, allowing for simplified design. By connecting the additional control elements with a new performance criterion, improved overall performance beyond the individual limitations of the subsystems is achieved.
Fig. 6. Cumulative subsystem servo error $e_r$ and $e_w$. Results show that the individual subsystem tracking performance is synchronized at low-frequencies by adding additional controller elements.

Fig. 7. Cumulative relative servo error $e_{wr}$. Results show a clear reduction in low-frequency content.

REFERENCES


