Supporting Information

■ SUPPLEMENT A. THE DETAILED DERIVATIONS OF THE COVARIANCE MATRIX OF THE RECONCILED ERRORS

To use a non-linear optimizer to find the values of the reconciled output $y_r(t)$ for the univariate system and $\mathbf{y}_r(t)$ for the multivariate system), just minimize the posterior distribution (of Eq.(25) for the univariate system and of Eq.(37) for the multivariate system). Both derivations are similar, so only the posterior distribution for the multivariate system is discussed. The partial derivatives of $p(\mathbf{y}_r(t)|\mathbf{y}_m(t),\hat{\mathbf{y}}(t))$ (Eq.(36)) with respect to $\mathbf{y}_r(t)$ can be expressed analytically using the equation

$$\frac{\partial p(\mathbf{y}_r(t)|\mathbf{y}_m(t),\hat{\mathbf{y}}(t))}{\partial \mathbf{y}_r(t)} \propto \mathbf{\Sigma}_{\boldsymbol{\varepsilon}}{}^{-1}(\mathbf{y}_r-\mathbf{y}_m)+\mathbf{\Sigma}_{\boldsymbol{\delta}}{}^{-1}(\mathbf{y}_r-\hat{\mathbf{y}}) \tag{S1}$$

By setting the derivative of $p(\mathbf{y}_r(t)|\mathbf{y}_m(t),\hat{\mathbf{y}}(t))$ with respect to $\mathbf{y}_r(t)$ to zero, after properly reorganizing the equation, the value of $\mathbf{y}_r(t)$ can be obtained,

$$\mathbf{y}_r=\hat{\mathbf{y}}+\mathbf{K}(\mathbf{y}_m-\hat{\mathbf{y}}) \tag{S2}$$

where $\mathbf{K}=(\mathbf{\Sigma}_{\boldsymbol{\varepsilon}}{}^{-1}+\mathbf{\Sigma}_{\boldsymbol{\delta}}{}^{-1})^{-1}\mathbf{\Sigma}_{\boldsymbol{\varepsilon}}{}^{-1}$. Thus the estimation error of the reconciled output ($\boldsymbol{\gamma}$) is

$$\begin{aligned}\boldsymbol{\gamma} &= \mathbf{y}_r - \mathbf{y} \\ &= \boldsymbol{\delta} + \mathbf{K}(\boldsymbol{\varepsilon} - \boldsymbol{\delta})\end{aligned} \tag{S3}$$

The covariance matrix of $\boldsymbol{\gamma}$ is

$$\begin{aligned}\mathbf{V}_r &= Cov(\boldsymbol{\gamma}) = Cov\left[\boldsymbol{\delta} + \mathbf{K}(\boldsymbol{\varepsilon} - \boldsymbol{\delta})\right] = Cov\left[\mathbf{K}\boldsymbol{\varepsilon} + (\mathbf{I} - \mathbf{K})\boldsymbol{\delta}\right] \\ &= \mathbf{K}^T \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} \mathbf{K} + (\mathbf{I} - \mathbf{K})^T \boldsymbol{\Sigma}_{\boldsymbol{\delta}} (\mathbf{I} - \mathbf{K})\end{aligned} \tag{S4}$$

where $\mathbf{K}^T \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} \mathbf{K}$ and $(\mathbf{I} - \mathbf{K})^T \boldsymbol{\Sigma}_{\boldsymbol{\delta}} (\mathbf{I} - \mathbf{K})$ can be rewritten as

$$\mathbf{K}^T \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} \mathbf{K} = \mathbf{K}^T \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} (\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1} + \boldsymbol{\Sigma}_{\boldsymbol{\delta}}^{-1})^{-1} \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1} = \mathbf{K}^T (\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1} + \boldsymbol{\Sigma}_{\boldsymbol{\delta}}^{-1})^{-1} \tag{S5}$$

$$\begin{aligned}(\mathbf{I} - \mathbf{K})^T \boldsymbol{\Sigma}_{\boldsymbol{\delta}} (\mathbf{I} - \mathbf{K}) &= \mathbf{K}^T (\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} \boldsymbol{\Sigma}_{\boldsymbol{\delta}}^{-1})^T \boldsymbol{\Sigma}_{\boldsymbol{\delta}} \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} \boldsymbol{\Sigma}_{\boldsymbol{\delta}}^{-1} \mathbf{K} \\ &= \mathbf{K}^T (\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} \boldsymbol{\Sigma}_{\boldsymbol{\delta}}^{-1})^T (\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1} + \boldsymbol{\Sigma}_{\boldsymbol{\delta}}^{-1})^{-1}\end{aligned} \tag{S6}$$

where the measurement noise are independent from each other; hence the covariance matrix of the measurement noise ($\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$) is diagonal.

After Eqs.(S5) and (S6) are substituted into Eq.(S4), the covariance matrix of $\boldsymbol{\gamma}$ can be obtained

$$\begin{aligned}\mathbf{V}_r &= \mathbf{K}^T (\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1} + \boldsymbol{\Sigma}_{\boldsymbol{\delta}}^{-1})^{-1} + \mathbf{K}^T (\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} \boldsymbol{\Sigma}_{\boldsymbol{\delta}}^{-1})^T (\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1} + \boldsymbol{\Sigma}_{\boldsymbol{\delta}}^{-1})^{-1} \\ &= [\mathbf{K}^T + \mathbf{K}^T (\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} \boldsymbol{\Sigma}_{\boldsymbol{\delta}}^{-1})^T](\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1} + \boldsymbol{\Sigma}_{\boldsymbol{\delta}}^{-1})^{-1} \\ &= \mathbf{K}^T [\mathbf{I} + (\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} \boldsymbol{\Sigma}_{\boldsymbol{\delta}}^{-1})^T](\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1} + \boldsymbol{\Sigma}_{\boldsymbol{\delta}}^{-1})^{-1} \\ &= \mathbf{K}^T \mathbf{K}^{-T} (\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1} + \boldsymbol{\Sigma}_{\boldsymbol{\delta}}^{-1})^{-1} = (\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1} + \boldsymbol{\Sigma}_{\boldsymbol{\delta}}^{-1})^{-1}\end{aligned} \tag{S7}$$

■ SUPPLEMENT B. DATA BASED ESTIMATION OF THE COVARIANCE MATRIX OF THE MODEL PREDICTION ERRORS.

Data based estimation of the covariance matrix ($\Sigma_{\boldsymbol{\delta}}$ for the multivariate system) of the model prediction errors would be derived as below. Because the derivations for the estimation of the variance η^2 for the univariate system and of the covariance $\Sigma_{\boldsymbol{\delta}}$ for the multivariate system are similar, so only the derivations for the multivariate system are discussed.

The multivariate system described (Eq.(7)) is operated by feedback-only controllers in the closed loop, and the controller is expressed in a transfer function form as

$$\mathbf{u}(t) = -\mathbf{C}(z^{-1})\mathbf{y}_{\mathbf{m}}(t) \tag{S8}$$

After substituting Eq.(S8) into Eq.(7), the corresponding measured outputs of the system for the disturbance changes and the measurement noise are

$$\mathbf{y}_m(t) = -\mathbf{A}(z^{-1})\mathbf{C}(z^{-1})\mathbf{y}_m(t) + \mathbf{B}(z^{-1})\mathbf{d}(t) + \boldsymbol{\varepsilon}(t) \tag{S9}$$

where $\mathbf{p}(t) = \mathbf{B}(z^{-1})\mathbf{d}(t) + \boldsymbol{\varepsilon}(t)$ is defined as the unmeasured part. Thus, the first term on the right side of Eq.(S9) can be estimated by a finite-length autoregressive model. Using the prediction error $\hat{\mathbf{p}}(t)$, the estimated covariance matrix of $\mathbf{p}(t)$ can be calculated by

$$\boldsymbol{\Sigma}_p = \operatorname{cov}\left[\hat{\mathbf{p}}(t)\right] = \operatorname{cov}[\mathbf{B}(z^{-1})\mathbf{d}(t) + \boldsymbol{\varepsilon}(t)] \tag{S10}$$

The disturbances to be considered in this paper are caused by the process equipment and only act on the process, the measurement noise is the variation in the sensor reading that does not correspond to changes in the process. Therefore, the disturbances and the measurement noise are assumed to be independent. the covariance matrix of $\mathbf{p}(t)$ is the linear combination of the covariance matrix of the model prediction error ($\mathbf{\Sigma_\delta}$) and the covariance matrix of the measurement noise ($\mathbf{\Sigma_\varepsilon}$),

$$\mathbf{\Sigma_p} = \mathrm{cov}[\mathbf{B}(z^{-1})\mathbf{d}(t)] + \mathrm{cov}[\boldsymbol{\varepsilon}(t)] = \mathbf{\Sigma_\delta} + \mathbf{\Sigma_\varepsilon} \tag{S11}$$

Thus, the estimated covariance matrix of the model prediction error can be calculated by

$$\mathbf{\Sigma_\delta} = \mathbf{\Sigma_p} - \mathbf{\Sigma_\varepsilon} \tag{S12}$$