I&EC research
Industrial & Engineering Chemistry Research

Article

Cite This: Ind. Eng. Chem. Res. 2019, 58, 2071−2083

pubs.acs.org/IECR

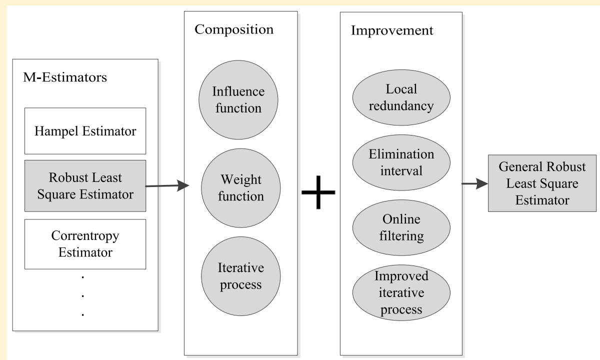
Novel Robust Least-Squares Estimator for Linear Dynamic Data Reconciliation

Rui Zhang, Yi Zhang, Yiping Feng, and Gang Rong*

State Key Laboratory of Industrial Control Technology, College of Control Science and Engineering, Zhejiang University, Hangzhou 310027, China

Supporting Information

ABSTRACT: Dynamic data reconciliation (DDR) is used to reduce the uncertainties in process measurement. Conventional data reconciliation theories and methods are based on least-squares estimation, whose conditions are hard to meet in real-world applications. Since least-squares estimators can be made robust by equivalent weight, many works concentrated on robust estimators and their performance in data reconciliation. However, robust least-squares algorithm is still unable to detect gross errors in low-redundancy variables. In this article, our work takes a step further in improving the robustness of the Huber estimator by taking structural redundancy into consideration. Modifications on influence function and weight function are made by involving local redundancy and rejection interval. The detectivity of gross error in low-redundant variables is increased. Combined with the above improvement and then taking advantage of online filtering, a novel robust dynamic data reconciliation algorithm is proposed. Case studies demonstrate the superiority of this methodology.



1. INTRODUCTION

Linear dynamic data reconciliation (LDDR) is a way of estimating variables on the basis of measurements and dynamic equilibrium equations. A dynamic material balance model can be represented by continuous-state space equations or a sampled input−output representation after discretization. Many other studies of LDDR have been carried out, but most of them did not properly deal with the complexity of dynamic systems and the demand for timeliness. Almasy (1990)[1] transformed the dynamic system balance problem into Kalman filtering problem by discretizing input and output variables. Darouach and Zasadzinski (1991)[2] transformed the LDDR into an algebraic problem by using the forward difference approximation method and obtained an iterative algorithm to avoid the truncation error and matrix singularity problem. The integral approach proposed by Bagajewicz and Jiang (1997)[3] provides another choice for solving the problem. However, it is difficult to separate a signal from noise. Bagajewicz (2000)[4] also applies this method to steady-state data reconciliation (SSDR) and DDR, proving that for linear systems without tank flow, the performance of both approaches is similar in the absence of biases and leaks.

Almasy (1990)[1] studied DDR based on Kalman filter and proved that Kalman filter equations are convenient for solving LDDR. Although Kalman filter has great characteristics in theory, it has to suppose that the state variables are self-correlated. To solve the general LDDR problem, the least-squares principles of linear SSDR for dynamic systems are extended. The new solution is based on the concept of generalized linear dynamic model, which was first introduced to data reconciliation (DR) by Darouach and Zasadzinski (1991).[2] Generalized linear dynamic model refers to the model formulation $Dx_{k+1} = Bx_k$, which contains more state variables than constraints. Because D is singular, the formulation cannot be written in a standard state space form. Obviously standard Kalman filter is not suitable for this situation, but we can utilize the recursive form of Kalman filter, which is more convenient for real-time processing. Bai (2006)[5] pointed out that DDR is actually a kind of filter technique. Based on the new model form, many studies have been working on a new recursive filtering algorithm. However, the model is quite huge, and the derivation procedure is complex and tedious.[6] Rollins and Devanathan (1993)[7] adopted the same model formulation and proposed a constrained least-squares estimator, making use of measurements at continuous adjacent time instants. Xu (2010)[8] proposed a new framework for generalized dynamic system, whose core part is a simplified least-squares formulation to represent the filtering problem.

The least-squares estimator (LSE) is the most widely used parameter estimation method. But the lack of robustness makes it imperfect in practice. Johnston and Kramer (1995)[9] reported the feasibility and better performance of robust estimators as the objective function in data reconciliation

Received: October 3, 2018
Revised: December 24, 2018
Accepted: December 30, 2018
Published: December 31, 2018

DOI: 10.1021/acs.iecr.8b04853
Ind. Eng. Chem. Res. 2019, 58, 2071−2083

especially when measurements are contaminated by gross errors. In the literature, DDR problems were addressed using fair function (FF),[10] three-part redescending estimator (TPRE),[11] correntropy (CO) function,[12] and so on. Subsequently, many studies have been conducted on robust estimators and their performance in data reconciliation. Including Arora and Biegler (2001),[13] many studies have demonstrated the potential of robust statistics developed by Huber (1981),[14] which attempts accurate estimation of statistical parameters. However, according to Özyurt (2004),[15]the influence function for weighted least-squares (WLS) is proportional to measurement error. The occurrence of gross errors will lead to incorrect estimations, so it is obviously not a robust estimator. But the research on quasi-weighted least-squares estimator (QWLS)[16] improved the robustness of WLS, which retained the advantage of WLS and, furthermore, enhanced the practicality in the presence of gross errors. A unified view on robust data reconciliation is provided using generalized objective functions within a probabilistic framework.[17] M-estimators are a maximum likelihood type estimator, which is one of the three main classes of robust estimators. Based on the principle of equivalent weights, M-estimator can be transformed into the form of classical least-squares. In other words, classical LSE can be made more robust by using equivalent weights.[18] This method not only improves the robustness of classical estimators but also preserves the advantages of simplicity in mathematical forms and computation procedure. Particularly, many well-developed mathematical models and computing methods based on least-squares are still applicable. The famous Huber estimator is the maximum likelihood estimation of the Huber function.[19] The Huber estimator is a kind of minimax estimation, which laid the foundation for location robust estimation. Huber[20] had made his first step toward attempting accurate estimation of statistical parameters in the presence of gross errors. Since that, robust estimators have been widely used in mathematical statistics, communication engineering, exploration mapping, and so on. From this point of view, data processing derived from Huber estimation in this paper is essentially a new robust least-squares (RLS) methodology, which is named general robust least-squares (GRLS).

After Huber, many researchers had paid much attention to improving other M-estimators. Ni (2009)[21] proposed two computational approaches for regression with stochastically bounded noise (RSBN). Hampel (2011)[22] focused on a smoothing principle for the Huber and other location M-estimators. Directed toward the correction of endogeneity problems in linear models, Kim (2007)[23] proposed a new robust estimator in the context of two-stage estimation methods. A new target function was taken into consideration based on robust M-estimators by Jin (2012).[24] Zhang (2010)[16] has proposed a quasi-weighted least-squares (QWLS) estimator to promote the performance of LSE. Claudia (2017)[25] distinguished systemic errors into outliers, biases, and drifts and also studied the performance of different M-estimators. However, few works have been carried out on the improvement of Huber estimator strategies for linear processes. The effectiveness of conventional algorithm depends on the reliability of approximated estimation solutions and reasonability of weight function during iterations. Conventional weight function applies variances with prior probability distributing measurement, disregarding the influence of robustness in structure space. Zhou had pointed out that the Huber estimator has no limitations on location of spatial distribution of observations.[18] In other words, gross errors may hardly be shown by residuals, but the introduction of local redundancy would raise the detectivity of gross errors especially for low-redundancy variables.

The novel GRLS estimator is proposed for LDDR, which has good performance in low-redundancy system and detecting gross errors. This algorithm focuses on making Huber estimator more robust based on the in-depth discussion on advantages and disadvantages of classical LSE. In this article, we improved the most important function of Huber estimation, the influence function, and weight function by considering construction redundancy. The algorithm procedure is combined with online filtering for performing online optimization easily. Due to the lack of accurate instruments and regular overhaul, low-redundant variables commonly exist, which has been a big challenge for DDR. The algorithm has good performance, especially for low-redundant variables. To avoid miscalculation introduced by the newly used local redundancy, the procedure has also been improved by combining constraints extracted from real process. The process has been tested in several cases and dozens of situations, not only listed in the article. The industrial case presented here is persuasive.

2. PROBLEM STATEMENT

In this section, we focus on improving Huber estimator. The main part of Huber distribution obeys normal distribution $N(0, \sigma^2)$, and the disturbance part obeys Laplace distribution.[18] The Huber estimator is the maximum likelihood estimation of the Huber function. Thus, the loss function of the Huber estimator is

$$\rho(v_i) = \begin{cases} \dfrac{v_i^2}{2\sigma_i^2} & |v_i| \le c\sigma_i \\ c\dfrac{|v_i|}{\sigma_i} - \dfrac{c^2}{2} & |v_i| > c\sigma_i \end{cases} \tag{1}$$

The influence function of Huber estimator is

$$\varphi(v_i) = \begin{cases} -c & v_i/\sigma_i < -c \\ v_i & -c \le v_i/\sigma_i \le c \\ c & v_i/\sigma_i > c \end{cases} \tag{2}$$

The weight function of Huber estimator is

$$w(v_i) = \begin{cases} 1 & |v_i/\sigma_i| \le c \\ c\sigma_i\,\mathrm{sign}(v_i)/v_i & |v_i/\sigma_i| > c \end{cases} \tag{3}$$

where v_i stands for residuals of the ith variables$v_i = \tilde{x}_i - \hat{x}_i$; z_i stands for measurements of the ith variable; $\hat{x}_i$ stands for reconciled values of the ith variable; σ_i stands for standard deviation of the ith variable; and c is a constant for Huber estimator, which is determined by trade-off between robustness and estimation efficiency. Estimation efficiency is the ratio of Cramer boundary and asymptotic variance. For Huber estimator, the efficiency is

$$\mathrm{eff} = \frac{1}{\mathrm{Var}} = \frac{[2\Phi(c) - 1]^2}{2\Phi(c) - 1 - 2c\phi(c) + 2c^2(1 - \Phi(c))} \tag{4}$$

DOI: 10.1021/acs.iecr.8b04853
Ind. Eng. Chem. Res. 2019, 58, 2071−2083

$\Phi(c)$ and $\phi(c)$ stand for distribution function and probability density value of standard normal distribution at point c. When $c = 1.4$, eff = 95.5%. In actual calculation, it should be determined by the rate of gross error δ:

$$\frac{1}{1-\delta} = 2\Phi(c) - 1 + 2\frac{\phi(c)}{c} \tag{5}$$

δ is always between 1% and 10%, so c is generally between 1 and 2. Mostly we choose c as around 1.5.

The effect of the Huber estimator is almost the same as LSE for tiny residuals. Thus, the estimator has high efficiency derived from Gaussian distribution. When the residuals are large, the performance of Huber estimator is approximately the same as least absolute deviations estimator. If gross errors appear, bounded weight will be assigned to the measurement biases to narrow the diffusion of gross error. However, for low-redundant observation, Huber estimator could not reflect its gross error by residual, which may lead to miscalculation. To preserve the advantage of Huber estimation, we propose a novel RLS estimator based on the improvement of the Huber one.

First, the influence function in eq 2 will be modified. Not only was the priori measurement variance replaced by measurement residual variance, but also the added local redundancy would be used to improve robustness in design space. The modified function will increase the detection rate of gross error, especially for low-redundancy variables. Second, the weight function eq 3 of Huber estimator will be reformulated to get rid of the contaminated measurements. Lastly, to weaken the disadvantage of introducing local redundancy, a concrete iterative flow of the novel robust LSE will be described in detail in section 4. The proposed improvements will be elaborated and proved to be effective in the rest of this article.

3. GENERAL FORMULATION

RLS estimator combines the robust estimation principle with the least-squares method. In the RLS estimation, a reasonable unbiased estimation of ρ function will be constructed to reflect the adopted scheme for abnormal measurements. This article uses the Huber function, which is not very sensitive to measurement biases. Some improvements based on the Huber function lead to better robustness. The general formulation of modeling and reconciliation is described in this section.

3.1. Local Redundancy. The intuitive concepts of observability and redundancy are explained in the literature. The concepts were first introduced by Kretsovalis and Mah (1988).[26] Carpentier et al. (1991)[27] split the variables into three classes according to redundancy, which is used to formulate a well-conditioned least-squares problem and to compute the values of redundant and calculable variables. Modron (1992)[28] offered a more complex classification of quantities. Maronna and Arcas (2009)[29] showed that the linear reconciliation problem can be represented by a standard multiple linear regression model, and the regression approach suggests a natural measure of the redundancy of an observation. Carpentier et al. (1991)[27] first defined a measure of the local redundancy, who called it error detectability of the ith measurement. Here we call it local redundancy in the following content. According to the footsteps, the concepts and methods used in this article are presented below.

The general formulation of data reconciliation is always described as eq 6. $\hat{X}$ and Z stand for reconciled variable vector and measured variable vector, respectively. $Q = \mathrm{diag}(\sigma_1, ..., \sigma_n)$ denotes the covariance matrix.

$$\begin{aligned} &\min(\hat{X} - Z)^{\mathrm{T}} Q^{-1}(\hat{X} - Z) \\ &\text{s.t.} \quad A\hat{X} = 0,\ A \in R^{m\times n},\ X \in R^n \end{aligned} \tag{6}$$

The measurement is generally described as follows.

$$z_i = x_i + \varepsilon_i, \quad i = 1, ..., n \tag{7}$$

where random noise $\varepsilon_i \sim N(0, \sigma_i)$ obeys normal distribution. The analytical solutions are listed below.

$$\hat{X} = (I - QA^{\mathrm{T}}(AQA^{\mathrm{T}})^{-1}A)Z \tag{8}$$

$$\Sigma_{\hat{x}} = \mathrm{cov}(\hat{x}) = (I - QA^{\mathrm{T}}(AQA^{\mathrm{T}})^{-1}A)Q \tag{9}$$

Given the following definition of sensitivity and redundancy respectively, H is redundancy.

$$M = I - QA^{\mathrm{T}}(AQA^{\mathrm{T}})^{-1}A \tag{10}$$

$$H = QA^{\mathrm{T}}(AQA^{\mathrm{T}})^{-1}A \tag{11}$$

We can obtain eq 12 and eq 13.

$$\hat{X} = MZ = (I - H)Z \tag{12}$$

$$Z - \hat{X} = HZ \tag{13}$$

Here, M reflects how sensitive the results are to the measurements. When Q is diagonal, the larger H_{ii} is, the smaller M_{ii} is. From eq 10, we know that the smaller the covariance matrix is, the more precise the estimator is.

When Q is diagonal, M and H have the following characteristics: M and H are idempotent matrices, $M^2 = M$, $H^2 = H$. $0 \le H_{ii} \le 1$, $0 \le M_{ii} \le 1$ $(i = 1, 2, ..., m)$, where H_{ii} is the diagonal element of H and M_{ii} is the diagonal element of M.

$$R_T = \mathrm{trace}(H) = \sum_{i=1}^{n} H_{ii} = m \tag{14}$$

$$\mathrm{trace}(M) = \sum_{i=1}^{n} M_{ii} = n - m \tag{15}$$

Here, R_T is defined as system redundancy by Bagajewicz (1999),[30] which is also called overall redundancy by Wang (2001).[31] H_{ii} is local redundancy which was first presented as comprehensive redundancy by Wang (2001).[31] However, they have been ignoring the influence of the precision of sensors on the system redundancy.

In the process of Figure 1, the precision of flowmeter of $S_i(i = 1, ..., 9)$ is defined as λ_i. Suppose that λ_3 is much higher than others in the measurement system; the measurement of S3 could also be calculated by (S_2, S_7) or (S_4, S_6). However, it is known that the precision of the calculated measurement must

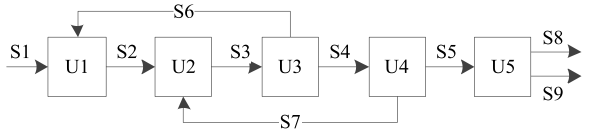


Figure 1. A simple process network.

DOI: 10.1021/acs.iecr.8b04853
Ind. Eng. Chem. Res. 2019, 58, 2071−2083

be lower than any component, so λ_3 would be much lower than its real value, which cannot play the role of reconciliation. Thus, the real precision of sensors should be concerned into redundancy calculation. From eq 11, H_{ii} involves both measurement precision Q and network structure A; then the system redundancy and measurement precision both have influence on comprehensive redundancy. H_{ii} is influential in the adjacent variables, which is also named as local redundancy.

It is noteworthy that the closer H_{ii} is to 0, the smaller the local redundancy of the variable is, and the greater its impact on the reconciliation accuracy is. When H_{ii} is equal to 0, X_i is not a redundant variable and cannot be reconciled.

3.2. Linear Dynamic Process Modeling. The general formulation of linear dynamic system is as follows:

$$\begin{cases} \frac{\mathrm{d}W}{\mathrm{d}t} = AF \\ CF = 0 \end{cases} \tag{16}$$

W and F denote the measurements of inventory and flow; A and C denote correlation coefficient matrix of capacity nodes and noncapacity nodes.

The measurement vector is as follows:

$$z_i = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}, i = 1, ..., n \tag{17}$$

In the form of Kalman filter, Z denotes measurement vector for all variables:

$$Z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} = H\begin{bmatrix} W \\ F \end{bmatrix} - v \tag{18}$$

Bagajewicz (1997)[3] pointed out specific variable classification methods for LDDR. According to the projection matrix transformation proposed by Crowe (1983),[32] unmeasured variables in eq 16 can be eliminated. Then the general formulation of the system is transformed into

$$\begin{cases} B_R\frac{\mathrm{d}W_R}{\mathrm{d}t} = A_RF_R \\ C_RF_R = 0 \end{cases} \tag{19}$$

$$Z_R = I\begin{bmatrix} W_R \\ F_R \end{bmatrix} - v_R \tag{20}$$

where all the matrices labeled with index R represent the obtained measured variables after matrix transformation.

After a series of transformations, the above formulations are transformed into the following difference equations, and the detailed process is described in the Supporting Information.

$$EX(i+1) = BX(i) \tag{21}$$

$$Z_R(i+1) = DX(i+1) - v_R(v+1) \tag{22}$$

where the coefficient matrix E, B, D are defined as follows:

$$E = [B_R \vdots - T_s(A_{R2} - A_{R1}C_{R1}^{-1}C_2)] \tag{23}$$

$$B = [B_R \vdots 0] \tag{24}$$

$$D = [I_1 \vdots I_{22} - I_{21}C_{R1}^{-1}C_{R2}] \tag{25}$$

where $C_R = [C_{R1} \vdots C_{R2}]$; note that C_{R1} is a square matrix and $|C_{R1}| \neq 0$; correspondingly $A_R = [A_{R1} \vdots A_{R2}]$; T_s is sampling period.

3.3. Robust Data Reconciliation. Xu and Rong (2010)[8] proposed a simplified least-squares formulation to represent the filtering problem in generalized linear dynamic systems. In our work, the least-squares estimator was replaced by the improved Huber estimator, and an iterative solution procedure is designed. To initialize the problem, it is considered that only random errors are present.

The generalized linear dynamic process model can be transformed into eq 21 and eq 22 by deleting unmeasured variables, eliminating algebraic equations, and discretization. Following the former section, the dynamic models will be combined with robust least-squares and filtering ideas.

Suppose the reconciliation results of time i were known as $\hat{X}(i|i)$, and its covariance matrix is $\Sigma(i|i)$, introducing auxiliary variables Y, $\tilde{Y}$, and ζ.

$$Y = BX(i) = EX(i+1) \tag{26}$$

$$\tilde{Y} = B\hat{X}(i|i) \tag{27}$$

$$\zeta = Y - \tilde{Y} \tag{28}$$

The covariance matrix of ζ is $B\Sigma(i|i)B^{\mathrm{T}}$.

Then, eq 21 and eq 22 can be transformed into eq 29.

$$\begin{bmatrix} \tilde{Y} \\ Z_R(i+1) \end{bmatrix} = \begin{bmatrix} E \\ D \end{bmatrix} X(i+1) - \begin{bmatrix} \zeta \\ v_R(i+1) \end{bmatrix} \tag{29}$$

In this form, dynamic data reconciliation model could be reconstructed based on robust least-squares theory.[18]

$$\min J = \sum_{j=1}^{n} Q_{jj}^{-1}\rho(v_{R,j}(i+1)) + \frac{1}{2}\zeta^{\mathrm{T}}[B\Sigma(i|i)B^{\mathrm{T}}]^{-1}\zeta$$
$$\text{s.t.} \quad \begin{bmatrix} B\hat{X}(i|i) \\ Z_R(i+1) \end{bmatrix} = \begin{bmatrix} E \\ D \end{bmatrix} X(i+1) - \begin{bmatrix} \zeta \\ v_R(i+1) \end{bmatrix} \tag{30}$$

where $Z_R(i+1)$ is measurement variable, $v_{R,j}(i+1)$ is the measurement bias of the jth variable on $(i+1)$th moment, Q is covariance matrix of measurements, Q_{jj} is the jth diagonal element of Q, and $\rho(v_{R,j}(i+1))$ is the loss function. The adopted Huber function is as follows:

$$\rho(v_{R,j}(i+1)) = \begin{cases} [v_{R,j}(i+1)]^2/2, \\ \quad |v_{R,j}(i+1)| \leq c\cdot\sqrt{Q_{jj}} \\ c\cdot|v_{R,j}(i+1)| - \frac{1}{2}c^2, \\ \quad |v_{R,j}(i+1)| > c\cdot\sqrt{Q_{jj}} \end{cases} \tag{31}$$

The solution of problem 30is obtained as follows. The derivative of eq 30 with respect to $X(I+1)$is

DOI: 10.1021/acs.iecr.8b04853
Ind. Eng. Chem. Res. 2019, 58, 2071−2083

$$\begin{bmatrix} E \\ D \end{bmatrix}^{T} \begin{bmatrix} [B\Sigma(i|i)B^{T}]^{-1} & 0 \\ 0 & \overline{P} \end{bmatrix} \begin{bmatrix} \zeta \\ v_{R}(i+1) \end{bmatrix} = 0 \tag{32}$$

where
$$\overline{P} = \begin{bmatrix} \overline{p}_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \overline{p}_n \end{bmatrix} \tag{33}$$

$$\overline{p}_j = \begin{cases} \dfrac{1}{Q_{jj}} & |v_{R,j}(i+1)| \le c\sqrt{Q_{jj}} \\ \dfrac{c\cdot \text{sign}[v_{R,j}(i+1)]}{Q_{jj}\cdot v_{R,j}(i+1)} & |v_{R,j}(i+1)| > c\sqrt{Q_{jj}} \end{cases} \tag{34}$$

The solution of eq 32 is

$$\hat{X}(i|i) = [E^{T}(B\Sigma(i-1|i-1)B^{T})^{-1}E + D^{T}Q^{-1}D]^{-1} \times [E^{T}(B\Sigma(i-1|i-1)B^{T})^{-1}B\hat{X}(i-1|i-1) + D^{T}Q^{-1}Z_{R}(i)] \tag{35}$$

To be similar to filtering formulations, the covariance matrix of $\hat{X}(i+1|i+1)$ should be iteratively calculated. Equation 30 is essentially a robust least-squares problem. According to the covariance matrix formulation of RLS estimator, the covariance matrix of $\hat{X}(i+1|i+1)$ is as follows.

$$\Sigma(i+1|i+1)= M^{-1}[E^{T}\vdots D^{T}] \begin{bmatrix} [B\Sigma(i|i)B^{T}]^{-1} & 0 \\ 0 & Q^{-1} \end{bmatrix} T \begin{bmatrix} [B\Sigma(i|i)B^{T}]^{-1} & 0 \\ 0 & Q^{-1} \end{bmatrix} \begin{bmatrix} E \\ D \end{bmatrix} M^{-1} \tag{36}$$

where

$$T = \begin{bmatrix} T_1 & 0 \\ 0 & T_2 \end{bmatrix} \tag{37}$$

$$T_1 = \begin{bmatrix} E\{\Phi'(\zeta_1)\}^2 & \cdots & E\{\Phi'(\zeta_1)\Phi'(\zeta_m)\} \\ \vdots & \ddots & \vdots \\ E\{\Phi'(\zeta_m)\Phi'(\zeta_1)\} & \cdots & E\{\Phi''(\zeta_m)\}^2 \end{bmatrix} \tag{38}$$

$$T_2 = \begin{bmatrix} E\{\rho'(v_{R,1})\}^2 & & \\ & \ddots & \\ & & E\{\rho'(v_{R,n})\}^2 \end{bmatrix} \tag{39}$$

$$M = [E^{T}\vdots D^{T}] \begin{bmatrix} G_1 & \\ & G_2 \end{bmatrix} \begin{bmatrix} E \\ D \end{bmatrix} \tag{40}$$

and

$$G_1 = \begin{bmatrix} [B\Sigma(i|i)B^{T}]^{-1}_{1,1}E\{\Phi''(\zeta_1)\} & \cdots & [B\Sigma(i|i)B^{T}]^{-1}_{1,m}E\{\Phi''(\zeta_m)\} \\ \vdots & \ddots & \vdots \\ [B\Sigma(i|i)B^{T}]^{-1}_{m,1}E\{\Phi''(\zeta_1)\} & \cdots & [B\Sigma(i|i)B^{T}]^{-1}_{m,m}E\{\Phi''(\zeta_m)\} \end{bmatrix} \tag{41}$$

$$G_2 = \begin{bmatrix} Q_{1,1}^{-1}E\{\rho''(v_{R,1})\} & & \\ & \ddots & \\ & & Q_{n,n}^{-1}E\{\rho''(v_{R,n})\} \end{bmatrix} \tag{42}$$

Here $\Phi(\cdot)$ uses $\zeta^2/2$, so

$$T_1 = B\Sigma(i|i)B^{T} \tag{43}$$

$$T_2 = \begin{bmatrix} E\{\rho'(v_{R,1})\}^2 & & \\ & \ddots & \\ & & E\{\rho'(v_{R,n})\}^2 \end{bmatrix} \tag{44}$$

S_1, S_2 are defined as follows:

$$S_1 = \begin{bmatrix} [B\Sigma(i|i)B^{T}]^{-1}_{1,1}E\{\Phi''(\zeta_1)\} & \cdots & [B\Sigma(i|i)B^{T}]^{-1}_{1,m}E\{\Phi''(\zeta_m)\} \\ \vdots & \ddots & \vdots \\ [B\Sigma(i|i)B^{T}]^{-1}_{m,1}E\{\Phi''(\zeta_1)\} & \cdots & [B\Sigma(i|i)B^{T}]^{-1}_{m,m}E\{\Phi''(\zeta_m)\} \end{bmatrix} = [B\Sigma(i|i)B^{T}]^{-1} \tag{45}$$

$$S_2 = \begin{bmatrix} Q_{1,1}^{-1}E\{\rho''(v_{R,1})\} & & \\ & \ddots & \\ & & Q_{n,n}^{-1}E\{\rho''(v_{R,n})\} \end{bmatrix} \tag{46}$$

So M is stated as

$$M = E^{T}[B\Sigma(i|i)B^{T}]^{-1}E + D^{T}S_2D \tag{47}$$

The covariance matrix becomes

$$\Sigma(i+1|i+1) = [E^{T}(B\Sigma(i|i)B^{T})^{-1}E + D^{T}S_2D]^{-1}[E^{T}(B\Sigma(i|i)B^{T})^{-1}E + D^{T}Q^{-1}T_2Q^{-1}D] \times [E^{T}(B\Sigma(i|i)B^{T})^{-1}E + D^{T}S_2D]^{-1} \tag{48}$$

As deduction above, the robust least-squares data reconciliation formulation is eq 35 and eq 48. The calculations of relevant parameters P, T_2, and S_2 are listed in eqs 33, 34, 44, and 46.

3.4. Procedure of Robust Data Reconciliation Using the Huber eEstimator. The DDR method eliminates the unmeasured variables by projection matrix and transforms it into a discrete model. We define T as the reconciliation period, and sampling frequency is presented by f. The algorithm

DOI: 10.1021/acs.iecr.8b04853
Ind. Eng. Chem. Res. 2019, 58, 2071−2083

process is given as following, and the flowchart of the algorithm is presented in Figure 2.

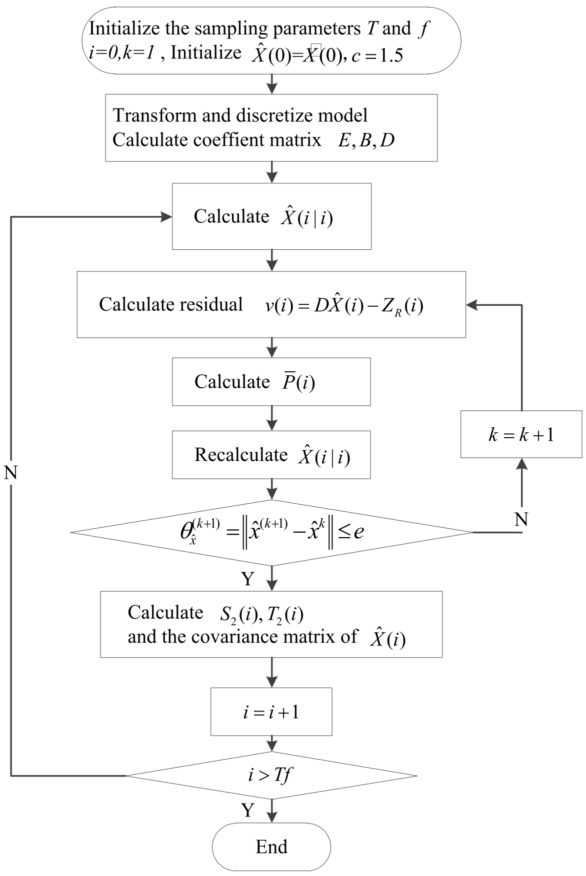


Figure 2. Flowchart of robust least-squares dynamic data reconciliation.

Step 1: Initialize the sampling parameters period T and frequency f, $i = 0$, $k = 1$, $\hat{X}(0) = \tilde{X}(0)$, $c = 1.5$;

Step 2: Transform and discretize the linear dynamic process model as presented in section 3.2; calculate coefficient matrix E, B, and D following eq 23, eq 24, and eq 25;

Step 3: Calculate the reconciled value $\hat{X}(i|i)$ according to eq 35;

Step 4: Calculate the residual $v(i) = D\hat{X}(i) - Z_R(i)$;

Step 5: Calculate the test statistics $\varepsilon_i = \frac{|v(i)|}{\sigma_i}$; if it is greater than the constant c, assign the weight according to eq 3;

Step 6: Calculate $\bar{P}(i)$ according to the obtained residual;

Step 7: Recalculate $\hat{X}(i|i)$;

Step 8: If the Euclidean distance $\theta_{\hat{x}}^{(k+1)} = \|\hat{x}^{(k+1)} - \hat{x}^k\|$ is smaller than a tiny constant e, go on to the next step; otherwise, $k = k + 1$ and go back to step 4;

Step 9: Calculate $S_2(i)$, $T_2(i)$, and the covariance matrix of $\hat{X}(i)$;

Step 10: $i = i + 1$, calculate the dynamic data on the next moment until exceeding given period.

4. SOLUTION STRATEGY

In this section, an improved algorithm is proposed to deal with disadvantages of RLS estimator by taking both observation and structural redundancy into account. The modified influence function and weight function can help us recognize lower redundant variables with gross errors. Meanwhile, the disadvantage from introducing local redundancy would be weakened by a redesigned iterative flow.

4.1. Improved Influence Function of the Huber Estimator.

Leverage is a measure of the distance between the independent variable values of an observation and the others. High-leverage points are observations made at extreme or outliers of the independent variables. These points are lack of neighboring observations, which means that the appropriate regression model will approach to that special point. If the errors are homoscedastic, an observation's leverage score determines the degree of noise in the model's misprediction of that observation. Conventional Huber estimator applied variances from prior probability distribution of measurement disregarding the importance of space robustness. It has no limitations on location of spatial distribution of observations, which leads to poor robustness in structure space. In this section, we introduced local redundancy to take spatial characteristics into consideration.

The definition of H has been given in eq 11. The relationship of system redundancy R_T and local redundancy H_{ii} has been defined in eq 14 $R_T = \text{strace}(H) = \sum_{i=1}^{n} H_{ii} = m$. Because H_{ii} value just takes part of system redundancy, it could not be the criterion of the redundancy of system. Here average local redundancy is taken as a threshold. If the local redundancy of variable is lower than threshold, it is supposed to be weakly redundant.

$$H_{m,ii} = \frac{1}{n}\text{trace}(H_{ii}) < \frac{m}{n} \tag{49}$$

For linear system, the least-squares analytical solution is listed in eq 8 and

$$v = Z - \hat{X} = QA^{\mathrm{T}}(AQA^{\mathrm{T}})^{-1}AZ = HZ \tag{50}$$

v stands for residuals, Z stands for measurements, $\hat{X}$ stands for reconciled values, the diagonal element of matrix Q is σ_i^2, and σ_i stands for standard deviation.

From eq 50,

$$v_i = \Delta x_i = H_{ii}\tilde{x}_i + \sum_{j\neq i}^{ii} H_{ii}\tilde{x}_j \tag{51}$$

H_{ii} is the local redundancy of the ith variable. If H_{ii} is close to zero, the ith variable is a weak redundancy variable. The gross error of the ith measurement could not be reflected from Δx_i, so the weak redundancy variables would hide the gross errors.

The local robustness properties of general penalized M-estimators are studied via the influence function. The influence function of Huber estimation is eq 2.

For variables in the $[-c, c]$ range, function values would be proportional to residual errors. Variables beyond this range are regarded as suspect data. If there are low-redundancy variables, the absolute value of residuals is much smaller than actual errors. The gross errors would hide themselves in the normal value section. Thus, the influence and weight functions should be reformulated and improved.

DOI: 10.1021/acs.iecr.8b04853
Ind. Eng. Chem. Res. 2019, 58, 2071−2083

In this paper, local redundancy is introduced into the calculation of residuals.

$$Q_{\Delta x} = QA^{\mathrm{T}}(AQA^{\mathrm{T}})^{-1}AQ = HQ \tag{52}$$

Thus, we let $\sigma_{r,i} = \sqrt{H_{ii}}\sigma_i$ replace σ_i. Since $0 \leq H_{ii} \leq 1$, $v_j/\sigma_{r,j} \geq v_i/\sigma_i$, and the lower the redundancy is, the smaller H_{ii} is, the larger $v_j/\sigma_{r,j}$ is. Then the influence function is formulated as

$$\varphi(v_i) = \begin{cases} -c\sqrt{H_{ii}} & v_i/\sqrt{H_{ii}}\sigma_i < -c \\ v_i & -c \leq v_i/\sqrt{H_{ii}}\sigma_i \leq c \\ c\sqrt{H_{ii}} & v_i/\sqrt{H_{ii}}\sigma_i > c \end{cases} \tag{53}$$

By analyzing eq 53, the purpose of the optimization is to shorten the linear part of former influence function.

4.2. Improved Weight Function of the Huber Estimator. The weight function is the key to robust least-squares. The improvement of the weight function should consider both robustness and efficiency. Observations would be divided into normal, available, and gross error. Based on the three types of observation, the interval should be divided into reservation, descending weight, and rejection intervals, respectively.

The efficiency and reliability of robust least-squares data reconciliation is mostly based on the descending weight interval of influence function $w(v_i)$. Furthermore, for reliable $w(v_i)$, the residuals should reflect measurement errors as much as possible, so local redundancy H_{ii}should be large. The reliability of initial iterative solution will be quite important. When $k = 0$, $x^{(0)} = \tilde{x}$, the initial value is just measurement value. The residual is zero, so the first circulation is just conventional least-squares estimation. The existence of gross error would lead to bias. Meanwhile, conventional Huber function has no rejection interval, which leads to worse robustness. Considering all the above, the weight function is revised as below.

$$w(v_i) = \begin{cases} 0 & \max\left(\left|Y_i = \dfrac{v_i}{\sigma_i\sqrt{H_{ii}}}\right|\right) \\ c\sigma_i\sqrt{H_{ii}}\,\mathrm{sign}(v_i)/v_i & Y_i = \left|\dfrac{v_i}{\sigma_i\sqrt{H_{ii}}}\right| > c \\ 1 & Y_i = \left|\dfrac{v_i}{\sigma_i\sqrt{H_{ii}}}\right| \leq c \end{cases} \tag{54}$$

In this weight function, the largest error statistics Y_i will be weighted zero in the rejection interval, which means that would be marked and eliminated.

In addition, for actual industrial process, measurement values usually have boundaries. These boundaries are always used as constraints in optimizations. Thus, according to the actual production, boundaries are also introduced to decide whether the eliminated gross error in the last circulation is reasonable. If not, the second largest value would replace the maximum to be weighted zero.

4.3. Procedure of Novel Robust Least-Squares Data Reconciliation. The procedure of novel robust least-squares data reconciliation is as follows (Figure 3):

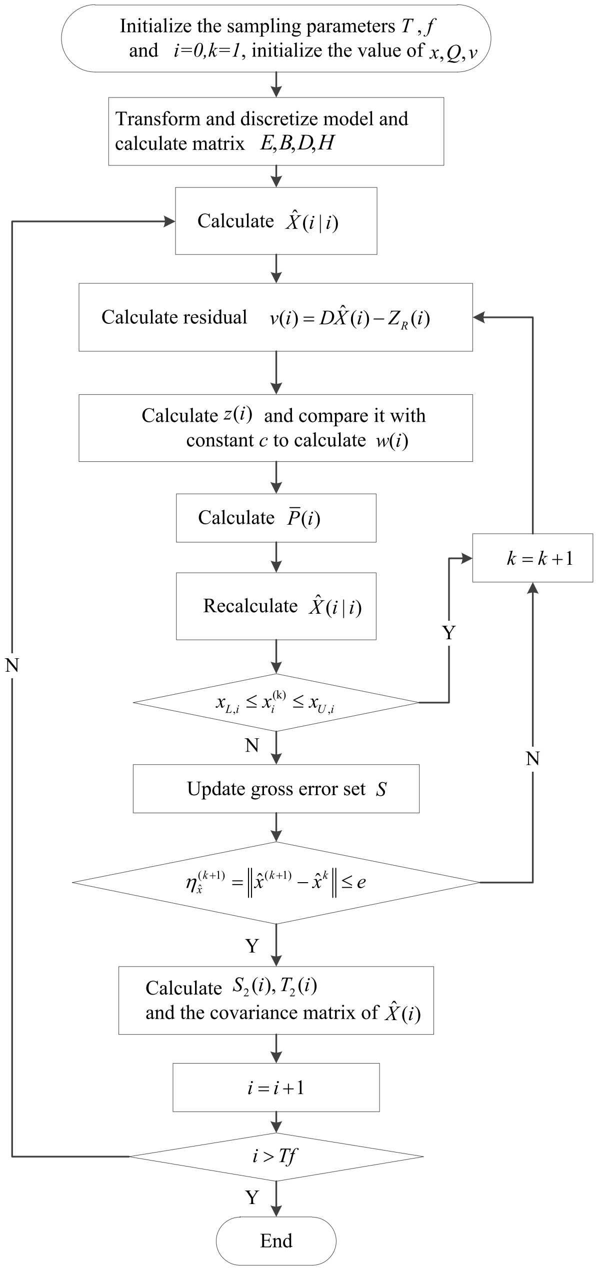


Figure 3. Flowchart of the new robust least-squares data reconciliation.

Step 1: Initialize the reconciliation period T and sampling frequency f, and $i = 0$, $k = 1$; When $k = 1$, the reconciled value $\hat{x}^{(1)} = (I - QA^{\mathrm{T}}(AQA^{\mathrm{T}})^{-1}A)z$, $Q = \mathrm{diag}(\sigma_i^2)$; residual $v^{(1)} = \hat{x}^{(1)} - z = -Hz$;

Step 2: Transform and then discretize the linear dynamic process model, calculate the coefficient matrix E, B, D, H;

Step 3: Calculate the reconciled value $\hat{X}(i|i)$ according to eq 35;

Step 4: Calculate the residual $v(i) = D\hat{X}(i) - Z_R(i)$;

Step 5: Calculate$\bar{P}(i)$ according to the obtained residual;

Step 6: Recalculate$\hat{X}(i|i)$;

Step 7: Calculate $Y_i = \frac{|v(i)|}{\sqrt{H_{ii}^{(k)}}\sigma_i}$; if it is greater than a constant c, then decrease its weight according to eq 54;

DOI: 10.1021/acs.iecr.8b04853
Ind. Eng. Chem. Res. 2019, 58, 2071−2083

Step 8: Check if the value satisfies $x_{L,i} \leq x_i^{(k)} \leq x_{U,i}$. If not, remove the latest added one in gross error set S and add the second suspicious one into S; otherwise, $k = k + 1$ and go back to step 4;

Step 9: If the Euclidean distance $\eta_{\hat{x}}^{(k+1)} = \|\hat{x}^{(k+1)} - \hat{x}^{k}\|$ is smaller than a tiny constant e, go on to the next step; otherwise, $k = k + 1$ and go back to step 4;

Step 10: Calculate $S_2(i)$, $T_2(i)$, and the covariance matrix of $\hat{X}(i)$;

Step 11: $i = i + 1$, calculate the dynamic data on the next moment until exceeding given period.

5. CASE STUDIES

The three methods, least-squares estimator, Huber estimator, and GRLS estimator, are applied to the same process. The following proposed performance indices are used to compare the effectiveness and robustness of the above three algorithms. The smaller the performance index η is, the better is the performance of the algorithm.

The sum of square error (SSE) is defined as below.

$$\mathrm{SSE} = \sum_{i=1}^{n} (x - \hat{x})^2 \tag{55}$$

The overall power (OP) is an index for correct detection rate of biased variable.

$$\mathrm{OP} = \frac{\text{number of bias identified correctly}}{\text{number of bias simulated}} \tag{56}$$

The average type I error (AVTI) is an index for wrong identification of a biased variable.

$$\mathrm{AVTI} = \frac{\text{number of bias indentified wrongly}}{\text{number of bias simulated}} \tag{57}$$

The average type II error (AVTII) is a measure for deviation failed to be identified.

$$\mathrm{AVTII} = \frac{\text{number of bias failed to be identified}}{\text{number of bias simulated}} \tag{58}$$

Accumulative reconciled bias:

$$\eta_1 = \sum_{i=1}^{n} (\hat{X}_i - X_{i,\mathrm{true}})^{\mathrm{T}} Q^{-1} (\hat{X}_i - X_{i,\mathrm{true}}) \tag{59}$$

Accumulative measured bias:

$$\eta_2 = \sum_{i=1}^{n} [Z_R(i) - X_{i,\mathrm{true}}]^{\mathrm{T}} Q^{-1} [Z_R(i) - X_{i,\mathrm{true}}] \tag{60}$$

Performance index:

$$\eta = \frac{\eta_1}{\eta_2} = \frac{\sum_{i=1}^{n} (\hat{X}_i - X_{i,\mathrm{true}})^{\mathrm{T}} Q^{-1} (\hat{X}_i - X_{i,\mathrm{true}})}{\sum_{i=1}^{n} [Z_R(i) - X_{i,\mathrm{true}}]^{\mathrm{T}} Q^{-1} [Z_R(i) - X_{i,\mathrm{true}}]} \tag{61}$$

5.1. Numerical Case.

The selected process is described in Figure 1, which was first introduced by Bagajewicz and Jiang.[3] It shows a simple process which contains nine flows and five nodes, among which U_1, U_2, U_3, and U_4 are capacity nodes and U_5 is noncapacity node. W and F represent volume and flow, respectively. Sampling interval is 1 s. Measurements are generated from actual data with normally distributed noise. To simplify the process of matrix transformation, suppose all the variables are measured. In the process of simulation, $c = 1.5$ for Huber function and tiny constant $e = 1e^{-6}$. Suppose the initial state are

$$W(0) = [100150100200]^{\mathrm{T}}$$

$$F(0) = [5152010510532]^{\mathrm{T}}$$

The standard deviation of each variable is 10% of the initial state.

The local redundancy of this system is

$$H_{ii} = [0.96640.97230.98540.98520.00240.0310 \cdots$$
$$0.02680.01230.00250.01210.003500]^{\mathrm{T}}$$

As stated in 3.1, when $H_{ii} = 0$, X_i is a nonredundant variable and could not be reconciled. So S8 and S9 are nonredundant. S1 and S5 have the greatest influence on the reconciliation of the system. The effect of H_{ii} in the robust reconciliation would be demonstrated afterward.

5.1.1. Scenario 1. In this scenario, we suppose that there is no gross error. Figure 4 shows the reconciled data of W_2 by both LS and GRLS. The two curves almost coincide. Table 1 also shows that the effects of both methods are almost the same when there is no gross error.

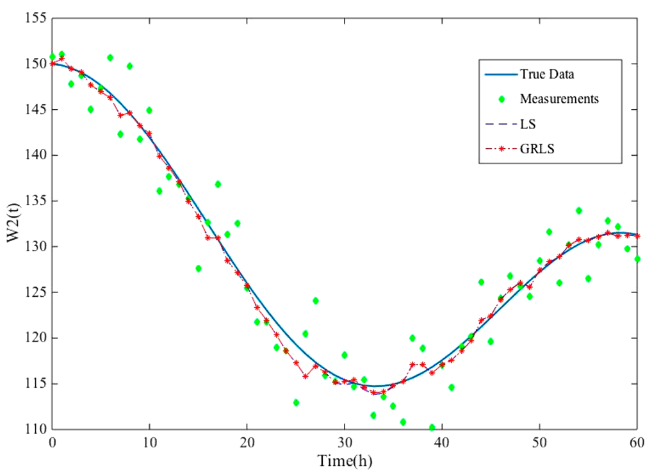


Figure 4. Reconciliation results comparison of W_2 (no gross error).

Table 1. Performance Comparison in Scenario 1

	η_1	η_2	η
LS	150.92	245.81	0.614
GRLS	148.74	245.81	0.605

5.1.2. Scenario 2. The process network is the same as presented in Figure 1. When the measurement of W_2 has gross errors from T5 to T7 and the magnitude of bias is 20, the reconciliation results are shown as Figure 5. The performance of conventional RLS are disappointing. Obviously, the GRLS shows much more robustness against gross error. We can see within three time-adjacent outliers, the GRLS algorithm has good performance as expected.

The performance indices of different methods are listed in Table 2; the GRLS shows better performance.

5.1.3. Scenario 3. When errors of W_1 begin from T45 and last at least 20 time cells, the amplitudes of the bias is still 20; the results are shown in Figure 6 and Figure 7. Both W_1 and

DOI: 10.1021/acs.iecr.8b04853
Ind. Eng. Chem. Res. 2019, 58, 2071−2083

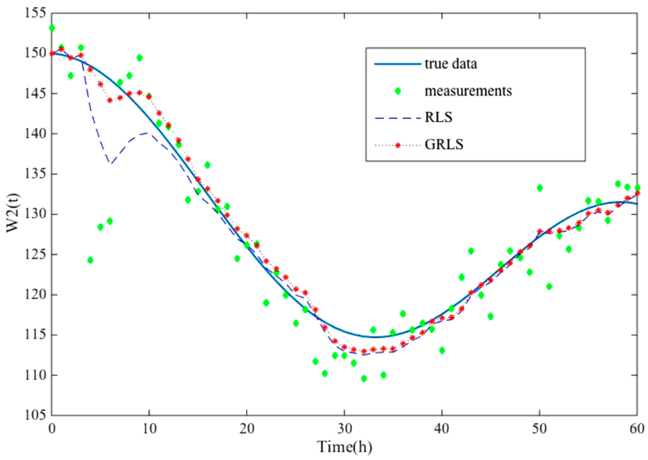


Figure 5. Reconciliation results comparison of W_2 (transient gross error).

Table 2. Performance Comparison in Scenario 2

	η_1	η_2	η
RLS	178.22	293.02	0.6082
GRLS	164.26	293.02	0.5606

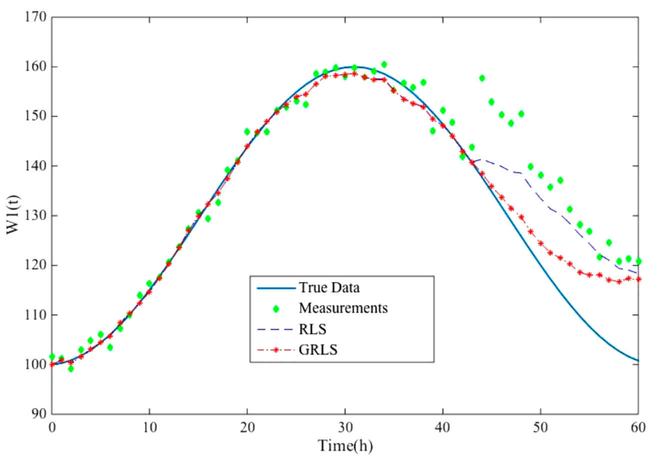


Figure 6. Reconciliation results comparison of W_1 (continuous gross errors).

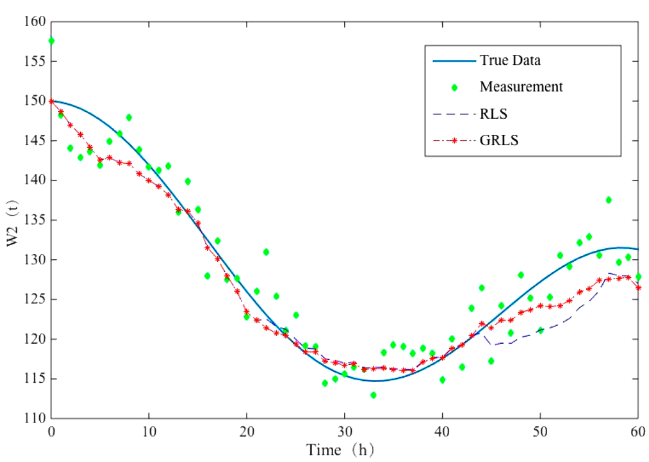


Figure 7. Reconciliation results comparison of W_2 (continuous gross errors).

W_2 are contaminated, and the validity of the traditional robust least-squares algorithm has been affected. Our improved algorithm shows better robustness which could also be revealed from Table 3. From the trend of line chart of

Table 3. Performance Comparison in Scenario 3

	η_1	η_2	η
RLS	457.25	783.79	0.5834
GRLS	260.74	783.79	0.3327

GRLS, almost after the gross error lasts over 10 h, the gap between GRLS and true data begins to be wider, but the gap between GRLS and RLS begins to be smaller. For W1 the line of GRLS coincides with RLS almost after 15 hours; for W2 the line of GRLS coincides with RLS almost after 17 hours. So the algorithm maybe break down when gross error lasts over 15 time units.

5.2. Comparative Case. A performance comparative analysis has been conducted. It includes not only redescending M-estimators like correntropy and adaptive Hampel but also some filtering methods. The process is the same as the above case in Figure 1.

5.2.1. Scenario 1. In this scenario, we suppose that there is no gross error. Figure 8 shows the reconciled bias of W_2 by the listed five methods, including correntropy, Kalman filter, GRLS, adaptive Hampel, and median filter. The results show that when there is no gross error, the four methods have almost the same performance, except for Kalman filter. The performance is judged by accumulative reconciled bias and listed in Table 4. The simulation time is 200 time units.

5.2.2. Scenario 2. In this scenario, we suppose that there are four discontinuous gross errors. The time and amplitudes of gross errors are (40, 45), (90, 25), (120, −35), and (180, −13). Figure 9 shows the reconciled bias of W_2 by the same five methods. The results show that when gross error occurs, the five methods show different performance. The accumulative reconciled biases are listed in Table 5. The Kalman filter has poor robustness with single gross error. Other methods had almost the same performance. It demonstrates that GRLS has good performance against gross error, and at least has the same robustness as correntropy and median filter.

5.2.3. Scenario 3. In this scenario, we suppose that there are continuous gross errors. First, deviation with amplitude around 25 and lasting for three time units is added to measurements of W_2. Figure 10a shows the reconciled bias of W_2 by five methods. The accumulative reconciled biases are listed in Table 6. When the disturbance lasts for just three time units, Kalman filter has performed worse, but others perform well.

When the disturbance lasts longer, adaptive Hampel and median filter meet their breakdown point as shown in Figure 10b, Table 7.

When the distance lasts much more longer, the robustness of adaptive Hampel is unstable. Hampel could not detect all the gross errors sometimes. But the tendency of GRLS is still stable as shown in Figure 10b and 10c, and the performance of GRLS is a little better than correntropy (Table 8).

Above all, most of the time, GRLS has better robustness than listed M-estimators and filters.

5.3. Industrial Case. The performance of the proposed GRLS algorithm has been tested using a classical industrial model first proposed by Serth (1986),[33] which was a methanol synthesis unit of a large chemical plant. The size and structure

DOI: 10.1021/acs.iecr.8b04853
Ind. Eng. Chem. Res. 2019, 58, 2071−2083

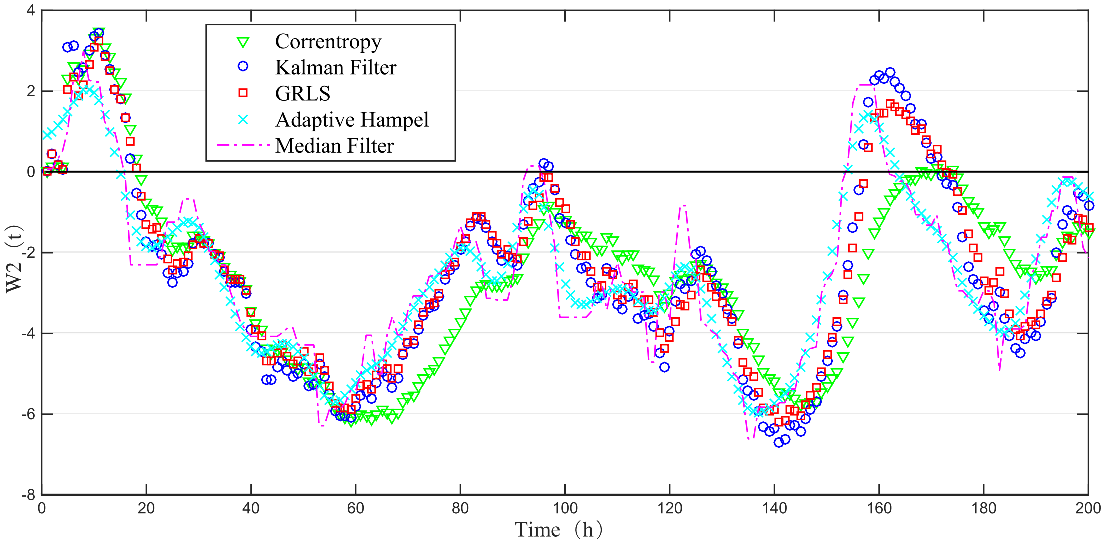


Figure 8. Reconciled bias of W_2 by five methods (no gross errors).

Table 4. Accumulative Reconciled Bias of Five Methods

	correntropy	Kalman	GRLS	adaptive Hampel	median
η_1	570.2244	604.1043	568.7501	553.5601	561.6709

Table 5. Accumulative Reconciled Bias of Five Methods

	correntropy	Kalman	GRLS	adaptive Hampel	median
η_1	568.5634	642.2739	566.1528	549.1144	566.9222

of this case are similar to many industrial systems. A network representation of the system consisting of 12 nodes and 28 streams is shown in Figure 11.

Assume that e_7, e_8, e_{11}, e_{26} are contaminated by 50% gross errors. True values, measurements, variances, redundancy, and local redundancy are listed, respectively. Reconciliation results of LS, RLS, and RLS-LR (robust least-squares considered local redundancy) and GRLS are compared.

The detailed results are shown in the Supporting Information. It can be concluded that the LSE algorithm has the worst robustness. Our novel robust least-squares iterative algorithm performs much better than other two. Traditional MT test based on LSE recognized 24 gross errors, but only 4 of them are right. Thus, the MT test based on LSE cannot locate the gross error correctly. After comparing the last three robust iterative algorithms, it is found that the RLS estimator detected that e_8, e_{26} contain gross errors, but e_7, e_{11} are missing. We found that the local redundancies of e_7, e_{11} are relatively lower, and the test statistics cannot reflect the gross error. The robust least-squares algorithm considering space robustness has detected gross errors in e_1, e_7, e_8, e_{11}, e_{22}, e_{26}, and e_{27}. Compared with the actual gross errors, there was no missing report, but the false report increased first kinds of errors. Because $0 \le H_{ii} < 1$, the introduction of H_{ii} would decrease the threshold of error detection and increase the detective rate of gross error, but at the same time, it improves the error statistics of other variables to some extent. The last column shows that

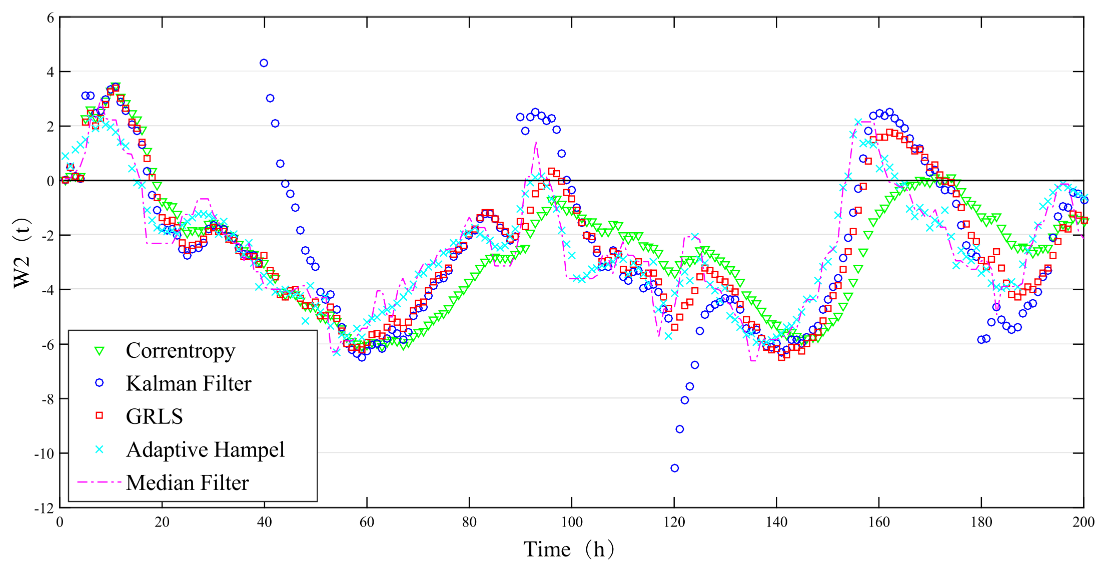


Figure 9. Reconciled bias of W_2 by five methods (discontinuous gross errors).

DOI: 10.1021/acs.iecr.8b04853
Ind. Eng. Chem. Res. 2019, 58, 2071−2083

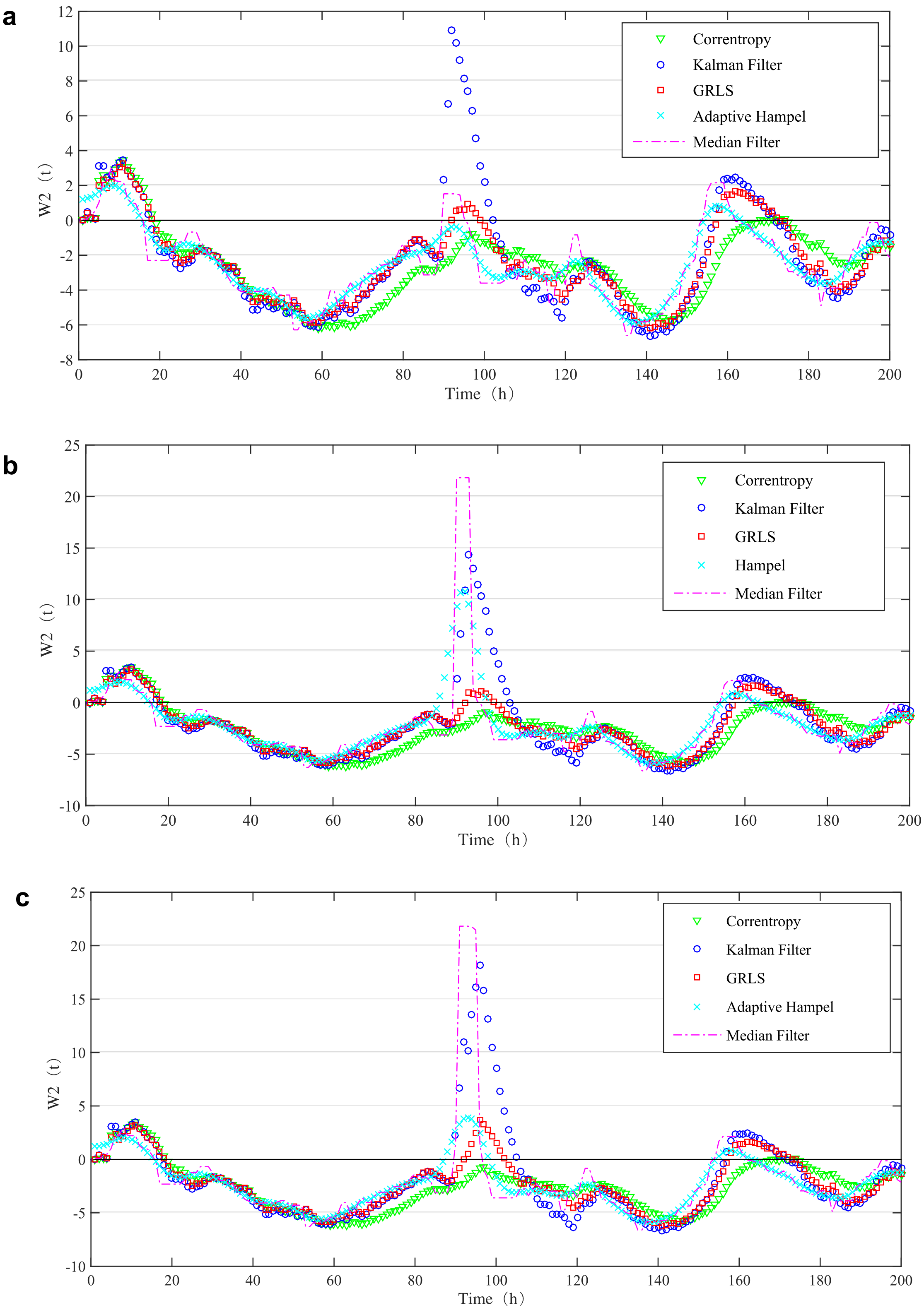


Figure 10. (a) Reconciled bias of W_2 (transient gross errors). (b) Reconciled bias of W_2 (gross errors last for four time cells). (c) Reconciled bias of W_2 (gross errors last for six time cells).

Table 6. Accumulative Reconciled Bias of Five Methods

	correntropy	Kalman	GRLS	adaptive Hampel	median
η_1	570.3373	672.9898	564.3034	541.8952	562.7660

Table 7. Accumulative Reconciled Bias of Five Methods

	correntropy	Kalman	GRLS	adaptive Hampel	median
η_1	573.4068	699.7748	565.4687	597.7854	643.9940

DOI: 10.1021/acs.iecr.8b04853
Ind. Eng. Chem. Res. 2019, 58, 2071−2083

Table 8. Accumulative Reconciled Bias of Five Methods

	correntropy	Kalman	GRLS	adaptive Hampel	median
η_1	577.1355	776.1355	572.8728	539.7410	716.4984

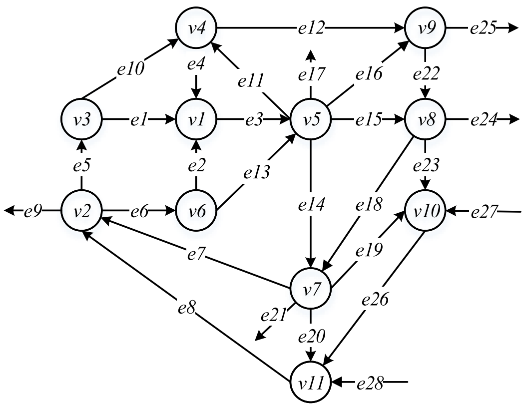


Figure 11. Network representation of steam system for a methanol synthesis unit.

our improved algorithm has the best performance. It located the four gross errors accurately.

After analyzing convergence, RLS iterated 15 times, RLS-LR iterated 36 times, and GRLS iterated 9 times; our algorithm costs the least computation resources and has good convergence.

To prove the statistical significance of the algorithm, the industrial case has been simulated 100 times. The average AVTI and AVTII are listed in Table 9. From the statistical results, since the conventional robust least-squares method has no consideration of space construction and redundancy, it has high possibility of misjudgment. Due to the introduction of local redundancy, RLS-LR has higher AVTI. After improving the algorithm flow, GRLS has added bounding constraints for variables into the iteration, and the failure rate of gross error detection is relatively much lower.

Table 9. AVTI and AVTII of Four Methods

	AVTI	AVTII
LS	83.33%	0.25%
RLS	0.13%	49.90%
RLS-LR	12.50%	0.08%
GRLS	0.083%	0.042%

6. CONCLUSION

In this work, we propose a novel robust least-squares estimator for linear DDR. The algorithm takes the structural robustness into consideration via using local redundancy in influence function and rejection interval in weight function. We redesign an iterative flow to weaken the disadvantage from introducing local redundancy. By taking advantage of the form of online filtering, and combining the with the improvement of Huber estimator, a novel robust dynamic data reconciliation algorithm is proposed. The novel estimator has almost the same effectiveness with RLS estimator without gross errors, but much better performance when the measurements are contaminated by gross errors. Compared to the conventional LSE or RLS estimator, the GRLS estimator improves the accuracy of DDR, and its application in industrial systems could improve the quality and accuracy of large-scale measurement. More challenges about implementing the method in complex industrial processes will be addressed in our future works.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.iecr.8b04853.

> Section 1: mathematical formulation: some formula deduction extended from section 3.2; section 2: data sheet of industrial case study, including comparison of different methods for methanol synthesis unit (Table 1) and comparison of different methods using SSE, OP, and AVTI (Table 2); section 3: nomenclature for important symbols (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*Tel.: +86 571 87953145. E-mail: grong@zju.edu.cn.

ORCID

Rui Zhang: 0000-0002-1023-9571
Yi Zhang: 0000-0002-0249-0547
Gang Rong: 0000-0002-6779-2968

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors gratefully acknowledge financial support by Smart Manufacturing Comprehensive Standardization and New Mode Application Project of MIIT (No.2017-ZJ-003) and Science Fund for Creative Research Groups of NSFC (Grant No. 61621002).

■ REFERENCES

(1) Almasy, G. A. Principles of dynamic balancing. *AIChE J.* **1990**, *36* (9), 1321−1330.

(2) Darouach, M.; Zasadzinski, M. Data reconciliation in generalized linear dynamic systems. *AIChE J.* **1991**, *37* (2), 193−201.

(3) Bagajewicz, M. J.; Jiang, Q. Integral approach to plant linear dynamic reconciliation. *AIChE J.* **1997**, *43* (10), 2546−2558.

(4) Bagajewicz, M. J.; Jiang, Q. Comparison of steady state and integral dynamic data reconciliation. *Comput. Chem. Eng.* **2000**, *24* (11), 2367−2383.

(5) Bai, S.; Thibault, J.; McLean, D. D. Dynamic data reconciliation: Alternative to Kalman filter. *J. Process Control* **2006**, *16* (5), 485−498.

(6) Veverka, V. V.; Madron, F. *Material and energy balancing in the process industries: From microscopic balances to large plants*; Elsevier: Amsterdam, Netherlands, 1997; Vol. 7.

(7) Rollins, D. K.; Devanathan, S. Unbiased estimation in dynamic data reconciliation. *AIChE J.* **1993**, *39* (8), 1330−1334.

(8) Xu, H.; Rong, G. A new framework for data reconciliation and measurement bias identification in generalized linear dynamic systems. *AIChE J.* **2010**, *56* (7), 1787−1800.

(9) Johnston, L. P.; Kramer, M. A. Maximum likelihood data rectification: Steady-state systems. *AIChE J.* **1995**, *41* (11), 2415−2426.

(10) (a) Fair, R. C. On the robust estimation of econometric models. In *Annals of Economic and Social Measurement*; NBER: Cambridge, MA, 1974; Vol. 3, number 4, pp 667−677. (b) Albuquerque, J. S.; Biegler, L. T. Data reconciliation and gross-error detection for dynamic systems. *AIChE J.* **1996**, *42* (10), 2841−2856.

(11) (a) Hampel, F. R. The influence curve and its role in robust estimation. *J. Am. Stat. Assoc.* **1974**, *69* (346), 383−393. (b) Nicholson, B.; López-Negrete, R.; Biegler, L. T. On-line state estimation of nonlinear dynamic systems with gross errors. *Comput. Chem. Eng.* **2014**, *70*, 149−159.
(12) (a) Chen, J.; Peng, Y.; Munoz, J. C. Correntropy estimator for data reconciliation. *Chem. Eng. Sci.* **2013**, *104*, 1019−1027. (b) Zhang, Z.; Chen, J. Correntropy based data reconciliation and gross error detection and identification for nonlinear dynamic processes. *Comput. Chem. Eng.* **2015**, *75*, 120−134.
(13) Arora, N.; Biegler, L. T. Redescending estimators for data reconciliation and parameter estimation. *Comput. Chem. Eng.* **2001**, *25* (11), 1585−1599.
(14) Huber, P. J. *Robust statistics*; John Wiley: New York, 1981.
(15) Özyurt, D. B.; Pike, R. W. Theory and practice of simultaneous data reconciliation and gross error detection for chemical processes. *Comput. Chem. Eng.* **2004**, *28* (3), 381−402.
(16) Zhang, Z.; Shao, Z.; Chen, X.; Wang, K.; Qian, J. Quasi-weighted least squares estimator for data reconciliation. *Comput. Chem. Eng.* **2010**, *34* (2), 154−162.
(17) Wang, D.; Romagnoli, J. A framework for robust data reconciliation based on a generalized objective function. *Ind. Eng. Chem. Res.* **2003**, *42* (13), 3075−3084.
(18) Zhou, J.-w.; Huang, Y.-c.; Yang, Y.-x. *Theory of Robustified Least Squares Method [M]*; The Press of Huazhong University of Science and Technology: Wuhan, China, 1997; Vol. *71* (95), pp 104−118.
(19) Huber, P. J. Robust estimation of a location parameter. *Ann. Math. Stat.* **1964**, *35* (1), 73−101.
(20) Huber, P. J. Robust statistics. In *International Encyclopedia of Statistical Science*; Springer: New York, 2011; pp 1248−1251.
(21) Ni, X. S.; Huo, X. Another look at Huber's estimator: A new minimax estimator in regression with stochastically bounded noise. *Journal of Statistical Planning and Inference* **2009**, *139* (2), 503−515.
(22) Hampel, F.; Hennig, C.; Ronchetti, E. A smoothing principle for the Huber and other location M-estimators. *Computational Statistics & Data Analysis* **2011**, *55* (1), 324−337.
(23) Kim, T.-H.; Muller, C. Two-stage Huber estimation. *Journal of Statistical Planning and Inference* **2007**, *137* (2), 405−418.
(24) Jin, S.; Li, X.; Huang, Z.; Liu, M. A New Target function for robust data reconciliation. *Ind. Eng. Chem. Res.* **2012**, *51* (30), 10220−10224.
(25) Llanos, C. E.; Sánchez, M. C.; Maronna, R. A. Classification of Systematic Measurement Errors within the Framework of Robust Data Reconciliation. *Ind. Eng. Chem. Res.* **2017**, *56* (34), 9617−9628.
(26) Kretsovalis, A.; Mah, R. Observability and redundancy classification in generalized process networks—I. Theorems. *Comput. Chem. Eng.* **1988**, *12* (7), 671−687.
(27) Carpentier, P.; Cohen, G. State estimation and leak detection in water distribution networks. *Civil Engineering Systems* **1991**, *8* (4), 247−257.
(28) Madron, F.; Veverka, V. Optimal selection of measuring points in complex plants by linear models. *AIChE J.* **1992**, *38* (2), 227−236.
(29) Maronna, R.; Arcas, J. Data reconciliation and gross error diagnosis based on regression. *Comput. Chem. Eng.* **2009**, *33* (1), 65−71.
(30) Bagajewicz, M. J.; Sanchez, M. C. Design and upgrade of nonredundant and redundant linear sensor networks. *AIChE J.* **1999**, *45* (9), 1927−1938.
(31) Xiuping, W.; Gang, R.; Shuqing, W. Key-Variable Analysis in Data Reconciliation. *Chinese J. Sci. Instrum.* **2001**, *22* (3), 235−239.
(32) Crowe, C.; Campos, Y.; Hrymak, A. Reconciliation of process flow rates by matrix projection. Part I: linear case. *AIChE J.* **1983**, *29* (6), 881−888.
(33) Serth, R.; Heenan, W. Gross error detection and data reconciliation in steam-metering systems. *AIChE J.* **1986**, *32* (5), 733−742.

DOI: 10.1021/acs.iecr.8b04853
Ind. Eng. Chem. Res. 2019, 58, 2071−2083