Supplementary material

Haoyu Wang, *Student Member, IEEE*, Zhengshuo Li, *Senior Member, IEEE*

I. Derivation process of the model based on the state transition equation

In this part the supplementary description about the derivation process of the state transition equation is given.

Based on reference [1], the state transition equation is derived as follows.

For a single pipeline, the dynamic changes in states, including the pressure π and mass flow rate q of gas flow, can be described by the following partial differential equations (PDEs):

$$\frac{\partial \pi}{\partial t}+\frac{c^2}{S}\frac{\partial q}{\partial x}=0 \tag{1}$$

$$\frac{\partial \pi}{\partial x}+\frac{1}{S}\frac{\partial q}{\partial t}+\frac{\lambda c^2}{2dS^2}\frac{q^2}{\pi}=0 \tag{2}$$

Then, we introduce the average flow velocity method to process the nonlinear term $\frac{\lambda c^2}{2dS^2}\frac{q^2}{\pi}$ in the equation (2).

Due to the following relationships between gas pressure, mass flow rate, flow density and flow velocity:

$$\pi=c^2\rho,\qquad q=\rho v S \tag{3}$$

So the nonlinear term in PDE (2) can be approximated as:

$$\frac{\lambda c^2}{2dS^2}\frac{(q_i^t)^2}{\pi_i^t}\approx\frac{\lambda c^2(\rho_i^t\bar{v}_i^t S)}{2dS^2c^2\rho_i^t}q_i^t=\frac{\lambda \bar{v}_i^t}{2dS}q_i^t \tag{4}$$

where $\bar{v}_i^t$ denotes the average flow velocity of the gas flow in pipe segment i during period t, which can be calculated by the following equation:

$$\bar{v}_i^t=\frac{2}{3}\frac{(v_{i+1}^t)^3-(v_i^t)^3}{(v_{i+1}^t)^2-(v_i^t)^2}=\frac{2}{3}\frac{c^2}{S}\frac{(q_{i+1}^t\pi_i^t)^2+q_{i+1}^t q_i^t\pi_{i+1}^t\pi_i^t+(\pi_{i+1}^t q_i^t)^2}{(\pi_{i+1}^t)^2\pi_i^t q_i^t+(q_i^t)^2\pi_{i+1}^t q_{i+1}^t} \tag{5}$$

By adopting the average flow velocity (4) and the Lax-Wendroff finite difference method, the PDEs (1) and (2) can be transformed into:

$$\left(\frac{\pi_{i+1}^{t+1}-\pi_{i+1}^t+\pi_i^{t+1}-\pi_i^t}{\Delta t}\right)+\frac{c^2}{S}\left(\frac{q_{i+1}^{t+1}-q_i^{t+1}+q_{i+1}^t-q_i^t}{\Delta x}\right)=0 \tag{6}$$

$$\begin{gathered}\left(\frac{\pi_{i+1}^{t+1}-\pi_i^{t+1}+\pi_{i+1}^t-\pi_i^t}{\Delta x}\right)+\frac{1}{S}\left(\frac{q_{i+1}^{t+1}-q_{i+1}^t+q_i^{t+1}-q_i^t}{\Delta t}\right)+\\ \frac{\lambda\bar{v}_i^t}{4dS}\left(q_{i+1}^{t+1}+q_{i+1}^t+q_i^{t+1}+q_i^t\right)=0\end{gathered} \tag{7}$$

To facilitate the expression of the matrix elements, we define the following coefficients:

$$\begin{cases} c_1=\dfrac{\pi^b}{\Delta t},c_2=\dfrac{c^2q^b}{S\Delta x},c_3=\dfrac{\pi^b}{\Delta x}, \\ d_1=q^b\left(\dfrac{1}{S\Delta t}+\dfrac{\lambda\bar{v}_i^t}{4dS}\right),d_2=q^b\left(\dfrac{1}{S\Delta t}-\dfrac{\lambda\bar{v}_i^t}{4dS}\right)\end{cases} \tag{8}$$

where π^b and q^b are the baseline values of the pressures and the mass flow rates, respectively.

Based on (1-8), the PDEs (1-6) and (1-7) of a discretized pipe segment i can be reformulated as follows:

$$\begin{gathered}\begin{bmatrix} c_1 & c_1 & -c_2 & c_2 \\ -c_3 & c_3 & d_1 & d_1\end{bmatrix}\left[\pi_i^{t+1}\quad \pi_{i+1}^{t+1}\quad q_i^{t+1}\quad q_{i+1}^{t+1}\right]^T= \\ \begin{bmatrix} c_1 & c_1 & c_2 & -c_2 \\ c_3 & -c_3 & d_2 & d_2\end{bmatrix}\left[\pi_i^t\quad \pi_{i+1}^t\quad q_i^t\quad q_{i+1}^t\right]^T\end{gathered} \tag{9}$$

where the pressures and the mass flow rates are per-unit values.

Based on (9), the discretized PDE matrix of a single pipeline can be formulated as follows:

$$\begin{gathered}\begin{bmatrix} c_1 & c_1 & & & & -c_2 & c_2 & & & \\ & c_1 & c_1 & & & & -c_2 & c_2 & & \\ & & \ddots & \ddots & & & & \ddots & \ddots & \\ & & & c_1 & c_1 & & & & -c_2 & c_2 \\ -c_3 & c_3 & & & & d_1 & d_1 & & & \\ & -c_3 & c_3 & & & & d_1 & d_1 & & \\ & & \ddots & \ddots & & & & \ddots & \ddots & \\ & & & -c_3 & c_3 & & & & d_1 & d_1\end{bmatrix}\begin{bmatrix}\boldsymbol{\pi}_l^{t+1}\\ \boldsymbol{q}_l^{t+1}\end{bmatrix}= \\ \begin{bmatrix} c_1 & c_1 & & & & c_2 & -c_2 & & & \\ & c_1 & c_1 & & & & c_2 & -c_2 & & \\ & & \ddots & \ddots & & & & \ddots & \ddots & \\ & & & c_1 & c_1 & & & & c_2 & -c_2 \\ c_3 & -c_3 & & & & d_2 & d_2 & & & \\ & c_3 & -c_3 & & & & d_2 & d_2 & & \\ & & \ddots & \ddots & & & & \ddots & \ddots & \\ & & & c_3 & -c_3 & & & & d_2 & d_2\end{bmatrix}\begin{bmatrix}\boldsymbol{\pi}_l^{t}\\ \boldsymbol{q}_l^{t}\end{bmatrix}\end{gathered} \tag{10}$$

$$\boldsymbol{C}_l^{t+1}\left[\boldsymbol{\pi}_l^{t+1}\quad \boldsymbol{q}_l^{t+1}\right]^T=\boldsymbol{C}_l^t\left[\boldsymbol{\pi}_l^t\quad \boldsymbol{q}_l^t\right]^T \tag{11}$$

where $\boldsymbol{y}_l^t=[\boldsymbol{\pi}_l^t\quad \boldsymbol{q}_l^t]^T$ denotes the vector of the pressures and mass flow rates in pipeline l at time t.

Equation (11) is a simplification of (10). Based on (11), the discretized PDE matrix for all pipelines in the natural gas system can be formulated as follows:

$$\begin{bmatrix}\boldsymbol{C}_1^{t+1} & & & \\ & \boldsymbol{C}_2^{t+1} & & \\ & & \ddots & \\ & & & \boldsymbol{C}_l^{t+1}\end{bmatrix}\boldsymbol{y}_{t+1}=\begin{bmatrix}\boldsymbol{C}_1^{t} & & & \\ & \boldsymbol{C}_2^{t} & & \\ & & \ddots & \\ & & & \boldsymbol{C}_l^{t}\end{bmatrix}\boldsymbol{y}_t \tag{12}$$

$$\boldsymbol{C}_{t+1}^{PN}\boldsymbol{y}_{t+1}=\boldsymbol{C}_t^{PN}\boldsymbol{y}_t \tag{13}$$

Equation (13) is a simplification of (12).

And there are the pressure continuity constraints, the flow balance constraints and compressor boost constraints at the junctions:

$$\pi_i^t=\pi_n^t,\forall i\in\Omega_n \tag{14}$$

$$\sum_{i\in\Omega_n^I}q_i^t-\sum_{j\in\Omega_n^O}q_j^t-q_n^t=0 \tag{15}$$

$$\pi_j^t=\pi_i^t+\pi_k^{CP},\forall i\in\Omega_k^I,j\in\Omega_k^O \tag{16}$$

where π_n^t denotes the pressure of practical node n at time t; Ω_n denotes the set of pressures π_i^t at the end of the pipeline connected to practical node n; q_n^t denotes the gas load of practical node n, which includes the traditional gas load $q_{n,t}^{GL}$ and the gas-fired unit gas withdrawal $q_{n,t}^{GF}$; Ω_n^I and Ω_n^O denote he sets of gas flows into and out of practical node n, respectively; π_k^{CP} denotes boosted pressure provided by compressor k; Ω_k^I and Ω_k^O denote he sets of gas flows into and out of compressor k.

Combined with these constraints, the dynamic state equation of the natural gas network can be formulated as follows:

$$\boldsymbol{C}_t^{NGS}[\boldsymbol{\pi}_t^{GW}\quad \boldsymbol{y}_t\quad \boldsymbol{q}_t^{GF}\quad \boldsymbol{q}_t^{GL}\quad \boldsymbol{\pi}_t^{CP}]^T=\boldsymbol{C}_{t-1}^{NGS}[\boldsymbol{\pi}_{t-1}^{GW}\quad \boldsymbol{y}_{t-1}]^T \tag{17}$$

where C_t^{NGS} denotes the coefficient matrix of the dynamic state equation of the natural gas system. Since different natural gas systems have different pressure continuity constraints, the flow balance constraints and compressor boost constraints, the parameters in C_t^{NGS} must be determined according to the specific topology of the natural gas system.

And with simple matrix operations, equation (17) can be transformed to:

$$\boldsymbol{O}(\boldsymbol{\lambda})\boldsymbol{y_t} = \boldsymbol{S}(\boldsymbol{\lambda})\boldsymbol{y_{t-1}} + \boldsymbol{I}\boldsymbol{u_t} \tag{18}$$

where $\boldsymbol{O}(\boldsymbol{\lambda})$ and $\boldsymbol{S}(\boldsymbol{\lambda})$ denotes the coefficient matrices containing the pipeline friction factors $\boldsymbol{\lambda}$ to be identified; $\boldsymbol{I}$ is a constant matrix; $\boldsymbol{u_t} = [\boldsymbol{\pi_{t-1}^{GW}} \quad \boldsymbol{\pi_t^{GW}} \quad \boldsymbol{\pi_t^{CP}} \quad \boldsymbol{q_t^{GL}} \quad \boldsymbol{q_t^{GF}}]^T$ denotes the control variables at time t.

Then through the elementary row transformation of the matrix, $\boldsymbol{y_t}$ can be decomposed into the state variables $\boldsymbol{y_t^{obs}}$ of the observable nodes and the state variables $\boldsymbol{y_t^{nob}}$ of the unobservable nodes, and the equation (18) can be transformed to:

$$\begin{bmatrix} \boldsymbol{y_t^{obs}} \\ \boldsymbol{y_t^{nob}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A_{11}} & \boldsymbol{A_{12}} \\ \boldsymbol{A_{21}} & \boldsymbol{A_{22}} \end{bmatrix} \begin{bmatrix} \boldsymbol{y_{t-1}^{obs}} \\ \boldsymbol{y_{t-1}^{nob}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{B_1} \\ \boldsymbol{B_2} \end{bmatrix} \boldsymbol{u_t} \tag{19}$$

Then through the recursive relationship and equation (19), we can get:

$$\begin{cases} \begin{bmatrix} \boldsymbol{y_t^{obs}} \\ \boldsymbol{y_t^{nob}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A_{11}} & \boldsymbol{A_{12}} \\ \boldsymbol{A_{21}} & \boldsymbol{A_{22}} \end{bmatrix} \begin{bmatrix} \boldsymbol{y_{t-1}^{obs}} \\ \boldsymbol{y_{t-1}^{nob}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{B_1} \\ \boldsymbol{B_2} \end{bmatrix} \boldsymbol{u_t} \\ \begin{bmatrix} \boldsymbol{y_{t-1}^{obs}} \\ \boldsymbol{y_{t-1}^{nob}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A_{11}} & \boldsymbol{A_{12}} \\ \boldsymbol{A_{21}} & \boldsymbol{A_{22}} \end{bmatrix} \begin{bmatrix} \boldsymbol{y_{t-2}^{obs}} \\ \boldsymbol{y_{t-2}^{nob}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{B_1} \\ \boldsymbol{B_2} \end{bmatrix} \boldsymbol{u_{t-1}} \\ \qquad \vdots \\ \begin{bmatrix} \boldsymbol{y_1^{obs}} \\ \boldsymbol{y_1^{nob}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A_{11}} & \boldsymbol{A_{12}} \\ \boldsymbol{A_{21}} & \boldsymbol{A_{22}} \end{bmatrix} \begin{bmatrix} \boldsymbol{y_0^{obs}} \\ \boldsymbol{y_0^{nob}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{B_1} \\ \boldsymbol{B_2} \end{bmatrix} \boldsymbol{u_1} \end{cases} \tag{20}$$

After simplification, equation (20) can be transformed to:

$$\begin{bmatrix} \boldsymbol{y_t^{obs}} \\ \boldsymbol{y_t^{nob}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A_{11}} & \boldsymbol{A_{12}} \\ \boldsymbol{A_{21}} & \boldsymbol{A_{22}} \end{bmatrix}^t \begin{bmatrix} \boldsymbol{y_0^{obs}} \\ \boldsymbol{y_0^{nob}} \end{bmatrix} + \sum_{\tau=0}^{t-1} \begin{bmatrix} \boldsymbol{A_{11}} & \boldsymbol{A_{12}} \\ \boldsymbol{A_{21}} & \boldsymbol{A_{22}} \end{bmatrix}^\tau \begin{bmatrix} \boldsymbol{B_1} \\ \boldsymbol{B_2} \end{bmatrix} \boldsymbol{u_{t-\tau}} \tag{21}$$

Then let $\boldsymbol{C_t} = \begin{bmatrix} \boldsymbol{C_{1,t}} & \boldsymbol{C_{2,t}} \\ \boldsymbol{C_{3,t}} & \boldsymbol{C_{4,t}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A_{11}} & \boldsymbol{A_{12}} \\ \boldsymbol{A_{21}} & \boldsymbol{A_{22}} \end{bmatrix}^t$ and $\boldsymbol{D_t} = \begin{bmatrix} \boldsymbol{D_{1,t}} \\ \boldsymbol{D_{2,t}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A_{11}} & \boldsymbol{A_{12}} \\ \boldsymbol{A_{21}} & \boldsymbol{A_{22}} \end{bmatrix}^t \begin{bmatrix} \boldsymbol{B_1} \\ \boldsymbol{B_2} \end{bmatrix}$, equation (21) can be transformed to:

$$\begin{bmatrix} \boldsymbol{y_t^{obs}} \\ \boldsymbol{y_t^{nob}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{C_{1,t}} & \boldsymbol{C_{2,t}} \\ \boldsymbol{C_{3,t}} & \boldsymbol{C_{4,t}} \end{bmatrix} \begin{bmatrix} \boldsymbol{y_0^{obs}} \\ \boldsymbol{y_0^{nob}} \end{bmatrix} + \sum_{\tau=0}^{t-1} \begin{bmatrix} \boldsymbol{D_{1,\tau}} \\ \boldsymbol{D_{2,\tau}} \end{bmatrix} \boldsymbol{u_{t-\tau}} \tag{22}$$

It is clear that from the equation (22) we can get:

$$\boldsymbol{y_t^{obs}} = \boldsymbol{C_{1,t}}\boldsymbol{y_0^{obs}} + \boldsymbol{C_{2,t}}\boldsymbol{y_0^{nob}} + \sum_{\tau=0}^{t-1} \boldsymbol{D_{1,\tau}}\boldsymbol{u_{t-\tau}} \tag{23}$$

Matrices C and D are mathematically related to matrices A and B as follows:

$$\begin{bmatrix} \boldsymbol{C_{1,t}} & \boldsymbol{C_{2,t}} \\ \boldsymbol{C_{3,t}} & \boldsymbol{C_{4,t}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A_{11}} & \boldsymbol{A_{12}} \\ \boldsymbol{A_{21}} & \boldsymbol{A_{22}} \end{bmatrix}^t, \qquad \begin{bmatrix} \boldsymbol{D_{1,t}} \\ \boldsymbol{D_{2,t}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A_{11}} & \boldsymbol{A_{12}} \\ \boldsymbol{A_{21}} & \boldsymbol{A_{22}} \end{bmatrix}^t \begin{bmatrix} \boldsymbol{B_1} \\ \boldsymbol{B_2} \end{bmatrix} \tag{24}$$

From this, we obtain two forms of state transition equations (19) and (23).

II. TEST RESULTS WITH DIFFERENT BASELINE VALUES AND HYPERPARAMETERS

TABLE I
MAPE OF THE PARAMETER ESTIMATION RESULTS WITH DIFFERENT BASELINE VALUES (UNIT: %)

π^b	q^b	MAPE
5×10^{6}	5×10^{1}	1.28×10^{-3}
5×10^{6}	1×10^{2}	1.18×10^{-3}
5×10^{6}	2×10^{2}	9.43×10^{-4}
1×10^{7}	5×10^{1}	6.60×10^{-4}
1×10^{7}	1×10^{2}	6.60×10^{-4}
1×10^{7}	2×10^{2}	6.63×10^{-4}
2×10^{7}	5×10^{1}	3.26×10^{-4}
2×10^{7}	1×10^{2}	3.26×10^{-3}
2×10^{7}	2×10^{2}	3.26×10^{-3}

TABLE II
MAPE OF THE PARAMETER ESTIMATION RESULTS WITH DIFFERENT BASELINE VALUES (UNIT: %)

System	γ	MAPE
10-node	30	1.05×10^{0}
	50	6.54×10^{-3}
	100	1.59×10^{-3}
	1000	6.63×10^{-4}
	2000	6.65×10^{-4}
20-node	30	5.77×10^{1}
	50	5.48×10^{-3}
	100	5.50×10^{-3}
	1000	5.59×10^{-3}
	2000	5.59×10^{-3}
40-node	30	1.05×10^{1}
	50	1.06×10^{-1}
	100	2.69×10^{-3}
	1000	3.66×10^{-3}
	2000	3.66×10^{-3}

III. THE PROOF OF THE OPTIMALITY OF THE CLOSED-FORM SOLUTION FOR THE FIRST STAGE

In the first stage of the manuscript, the goal is to estimate the initial value of unobservable state variables $\boldsymbol{y_0^{nob,est}}$. The optimization problem is defined by:

$$\min_{\boldsymbol{y_0^{nob,est}}} \sum_{t=1}^{K} \left\| \boldsymbol{y_t^{obs}} - \boldsymbol{C_{1,t}}\boldsymbol{y_0^{obs}} - \boldsymbol{C_{2,t}}\boldsymbol{y_0^{nob}} - \sum_{\tau=0}^{t-1} \boldsymbol{D_\tau}\boldsymbol{u_{t-\tau}} \right\|^2 \tag{25}$$

This is an unconstrained linear least-squares problem, equivalent to:

$$\min_{\boldsymbol{y}} \|\boldsymbol{A}\boldsymbol{y} - \boldsymbol{b}\|^2 \tag{26}$$

where $\boldsymbol{A} = [\boldsymbol{C_{2,1}} \quad \boldsymbol{C_{2,2}} \quad \cdots \quad \boldsymbol{C_{2,K}}]^T$ s the design matrix (dimension $K \times n_{nob}$, n_{nob} is the dimension of unobservable states); $\boldsymbol{b} = [\boldsymbol{r_1^{obs}} \quad \boldsymbol{r_2^{obs}} \quad \cdots \quad \boldsymbol{r_K^{obs}}]^T$ is the residual vector, $\boldsymbol{r_t^{obs}} = \boldsymbol{y_t^{obs}} - \boldsymbol{C_{1,t}}\boldsymbol{y_0^{obs}} - \sum_{\tau=0}^{t-1} \boldsymbol{D_\tau}\boldsymbol{u_{t-\tau}}$; $\boldsymbol{y} = \boldsymbol{y_0^{nob,est}}$ is the optimization variable.

Then the closed-form solution is given by:

$$\boldsymbol{y_0^{nob,est}} = \boldsymbol{A}^\dagger \boldsymbol{b} \tag{27}$$

where $\boldsymbol{A}^\dagger$ is the Moore-Penrose pseudo-inverse.

Since the optimization problem is unconstrained (no inequality or equality constraints), the KKT conditions reduce to the first-order optimality condition: the gradient of the objective function must be zero. For convex problems (e.g., least-squares), this condition is also sufficient, ensuring global optimality.

The objective function is:

$$f(\boldsymbol{y}) = \|\boldsymbol{A}\boldsymbol{y} - \boldsymbol{b}\|^2 = (\boldsymbol{A}\boldsymbol{y} - \boldsymbol{b})^T(\boldsymbol{A}\boldsymbol{y} - \boldsymbol{b}) \tag{28}$$

The gradient is:

$$\nabla f(\boldsymbol{y}) = 2\boldsymbol{A^T}(\boldsymbol{A}\boldsymbol{y} - \boldsymbol{b}) \tag{29}$$

KKT conditions (unconstrained simplification) require:

$$\nabla f(\boldsymbol{y}) = 0 \Rightarrow \boldsymbol{A^T}(\boldsymbol{A}\boldsymbol{y} - \boldsymbol{b}) = 0 \tag{30}$$

This is equivalent to the normal equation:

$$\boldsymbol{A^T}\boldsymbol{A}\boldsymbol{y} = \boldsymbol{A^T}\boldsymbol{b} \tag{31}$$

We now prove that the closed-form solution $\boldsymbol{y}_{\mathbf{0}}^{\boldsymbol{nob,est}} = \boldsymbol{A}^{\dagger}\boldsymbol{b}$ satisfies this condition:

Substituting $\boldsymbol{y}_{\mathbf{0}}^{\boldsymbol{nob,est}} = \boldsymbol{A}^{\dagger}\boldsymbol{b}$ into the normal equation:

$$\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}\boldsymbol{y}_{\mathbf{0}}^{\boldsymbol{nob,est}} - \boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{b} = (\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}\boldsymbol{A}^{\dagger} - \boldsymbol{A}^{\boldsymbol{T}})\boldsymbol{b} \tag{32}$$

Key Step: Prove $\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}\boldsymbol{A}^{\dagger} = \boldsymbol{A}^{\boldsymbol{T}}$.

Moore-Penrose pseudo-inverse properties:

1. $\boldsymbol{A}\boldsymbol{A}^{\dagger}\boldsymbol{A} = \boldsymbol{A}$
2. $\boldsymbol{A}^{\dagger}\boldsymbol{A}\boldsymbol{A}^{\dagger} = \boldsymbol{A}^{\dagger}$
3. $(\boldsymbol{A}\boldsymbol{A}^{\dagger})^{T} = \boldsymbol{A}\boldsymbol{A}^{\dagger}$
4. $(\boldsymbol{A}^{\dagger}\boldsymbol{A})^{T} = \boldsymbol{A}^{\dagger}\boldsymbol{A}$

Derivation of $\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}\boldsymbol{A}^{\dagger} = \boldsymbol{A}^{\boldsymbol{T}}$:

$$\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}\boldsymbol{A}^{\dagger} = \boldsymbol{A}^{\boldsymbol{T}}(\boldsymbol{A}\boldsymbol{A}^{\dagger}) = \boldsymbol{A}^{\boldsymbol{T}}(\boldsymbol{A}\boldsymbol{A}^{\dagger})^{T} = (\boldsymbol{A}(\boldsymbol{A}^{\dagger})^{T})\boldsymbol{A} \tag{33}$$

By property 3 $(\boldsymbol{A}\boldsymbol{A}^{\dagger})^{T} = \boldsymbol{A}\boldsymbol{A}^{\dagger}$:

$$\boldsymbol{A}^{\boldsymbol{T}}(\boldsymbol{A}\boldsymbol{A}^{\dagger}) = \boldsymbol{A}^{\boldsymbol{T}}(\boldsymbol{A}\boldsymbol{A}^{\dagger})^{T} = \left((\boldsymbol{A}\boldsymbol{A}^{\dagger})\boldsymbol{A}\right)^{T} = (\boldsymbol{A}\boldsymbol{A}^{\dagger}\boldsymbol{A})^{T} \tag{34}$$

By property 1 $\boldsymbol{A}\boldsymbol{A}^{\dagger}\boldsymbol{A} = \boldsymbol{A}$:

$$(\boldsymbol{A}\boldsymbol{A}^{\dagger}\boldsymbol{A})^{T} = \boldsymbol{A}^{\boldsymbol{T}} \tag{35}$$

So we can get:

$$\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}\boldsymbol{A}^{\dagger} = \boldsymbol{A}^{\boldsymbol{T}} \tag{36}$$

Substitute $\boldsymbol{y}_{\mathbf{0}}^{\boldsymbol{nob,est}} = \boldsymbol{A}^{\dagger}\boldsymbol{b}$ and $\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}\boldsymbol{A}^{\dagger} = \boldsymbol{A}^{\boldsymbol{T}}$:

$$\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}\boldsymbol{y}_{0}^{\boldsymbol{nob,est}} - \boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{b} = (\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}\boldsymbol{A}^{\dagger} - \boldsymbol{A}^{\boldsymbol{T}})\boldsymbol{b} = \mathbf{0} \tag{37}$$

Thus:

$$\nabla f\left(\boldsymbol{y}_{0}^{\boldsymbol{nob,est}}\right) = 0 \tag{38}$$

Since $f\left(\boldsymbol{y}_{0}^{\boldsymbol{nob,est}}\right)$ is convex (quadratic), $\boldsymbol{y}_{\mathbf{0}}^{\boldsymbol{nob,est}}$ is the global optimum.

References

[1] P. Zhao, Z. Li, X. Bai, J. Su, and X. Chang, “Stochastic real-time dispatch considering AGC and electric-gas dynamic interaction: Fine-grained modeling and noniterative decentralized solutions,” Applied Energy, vol. 375, p. 123976, Dec. 2024.