Modelling of high-speed railway traction power supply systems based on subspace identification

D Y Wu[1], L Ma[1] and S K Cheng[1]

[1] School of Electrical Engineering, Southwest Jiaotong University, Chengdu, 611756, China

malei@swjtu.edu.cn

Abstract. In this paper, reduced mathematical models for high-speed railway traction power systems based on subspace identification method (SIM) are proposed. SIM applies reliable algebra tools to estimate state variables, and then provides an analytic representation for the identified systems, especially those of large dimension. Accuracy for identification results of different orders is discussed, under both normal operation and fault cases. With appropriate accuracy-simplicity trade-off, a uniform order decision is deduced. Reduced models derived from identification could be used for further studies of network-side impact on train-network coupling problems.

1. Introduction

Thanks to the remarkable improvement of power electronic technology, high-speed trains with AC-DC-AC power drive system have become prevailing in rail transit arena, which largely promotes railway transportation capacity. Meanwhile, many train-network coupling problems have emerged, endangering the safety of railway operation. Typical phenomena include low-frequency oscillation (LFO), harmonic resonance and harmonic instability.[1]

Many scholars attribute these problems to AC-DC rectifiers in high-speed trains and focus coupling problem analysis on them. Traction network, a large distributed-parameter system with multiple lines in parallel, is usually simplified as a single invariant resistance.[1,2,3] However, the simplification does not consider the different states on the network side. A state-space model of traction power supply system based upon system topology has been proposed, but high-order structures add to computational and analytical complexity. [4]

This paper analyses the limitation of two mechanism-based traction network models. Considering the advantage of system identification, reduced mathematical models of traction network are constructed by subspace identification method (SIM). A uniform order decision for identification under normal operation and fault cases is reached. With variation of network models seen as uncertainty to trains, robustness of corresponding control designs for rectifiers can be improved.

2. Mechanism-based models of traction network

The all-parallel autotransformer-fed traction power supply system is a symmetrical linear multi-port network, shown in figure 1. It contains up to 14 lines: up track and down track messenger wires, contact wires, feeders, steel rails, protection wires and earth wires. Some of these conductors are parallel-connected continuously, thus electrically equivalent to one conductor line.[3,5] Commonly used patterns include 6-conductor, 8-conductor and 10-conductor models.

Two approaches for the traction network modelling have been widely used: simulation model and state-space model based on topology. Both of them derive from electrical connections and structures of traction power supply systems.

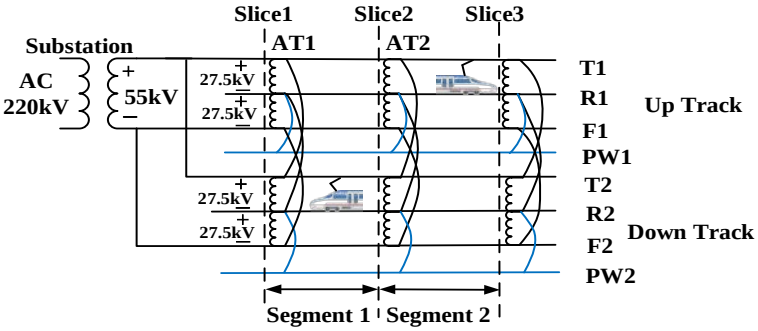


Figure 1. The typical 8-conductor model of traction network combines the earth wires and protection wires as the synthetized protection wires (PW1, PW2)

2.1 Simulation model

The simulation model represents distributed-parameter characteristics of traction network by dividing long wires into small segments. For each segment, a π-equivalent circuit is obtained from phase-modal conversion.[1] Thus the entire traction network is equal to a repetitive combination of π-equivalent circuits with impedance and admittance matrices, shown in figure 2.

The simulation model can show the transient responses and different states of train-network system. But it can't be used for further intrinsic mechanism analysis as simulation components are packaged.

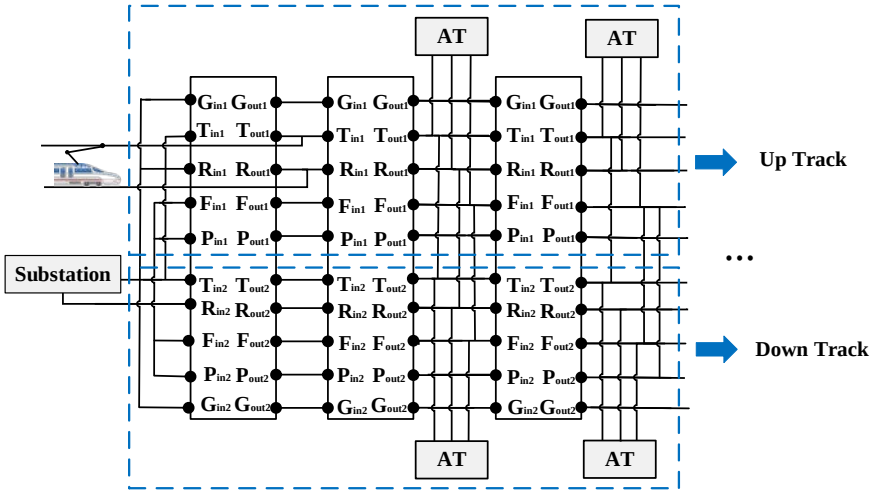


Figure 2. The simulation model of the traction network is equal to a chain circuit, with equivalent circuits of segments packaged as subsystems

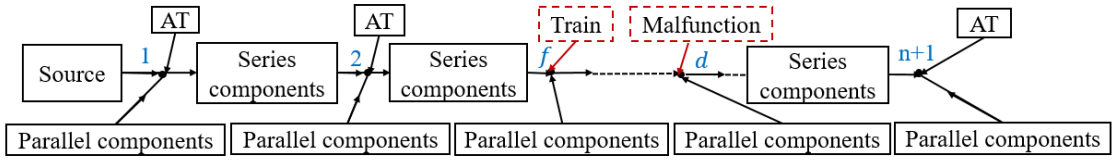
2.2 State-space model

[4] introduces a state-space model of the traction power supply system based on topology. Multiple lines of traction network are split into small segments, and impedances and mutual inductances of segments are seen respectively as series components and parallel elements joint at slicing points. Hence the traction power supply system is equal to a chain of modularized models: substation, AT, series components and parallel components, shown in figure 3.

At the train's location, a new section is created with train's current injected to network as new inputs. Typical wire faults create new nodes at their places. A very small resistance is added at the node where short-circuit happens, while a very large impedance is joint at the node to represent wire break. State functions come from the partial differential equations of dynamic elements in modules.

Although this state-space model can reflect the dynamic characteristics of network, it still has some limitation: 1) Due to segments and multiple electrical components of traction network, the state-space models are usually of very high orders even beyond 100, which adds up the computational complexity and difficulty for analysis. 2) Once train's location changes, or any fault of wires or AT occurs, the

topology of the traction network changes, with new sections created. New state-space models, reproduced by adjusting its corresponding equation in system matrices and state variables, are usually of different orders.



State variables: $i_s\ i_{AT1}\ i_{11}\ \cdots\ i_{16}\ u_{11}\ \cdots u_{16}\ \ i_{AT2}\ \ \cdots\ \ \cdots\ \ \cdots\ i_{train}\ \cdots\ \ i_{malfunction}\ \cdots$

Figure 3. Modularized diagram of traction network system

3. Identification method

3.1 Identification modelling

In practical application, some analytic models based on systems' physical laws seem too complicated, especially when studies focus on system's external features. In train-network coupling system, we mainly concern about the electrical variables at the joint point of train and network, so other state variables can actually be omitted.

Consequently, identification is applied to give a close description of complicated systems. With certain experiments, the system model can be conducted from collected input-output data and predefined hypotheses (like user-specified parametrization).[6] For given inputs, the identification result is supposed to provide outputs very close to collected data from original system. The uncertainty from accuracy impairment compared to mechanism-based mathematical models is allowed as long as the robustness of the overall system is ensured. In cases of controller designs for complex systems, system identification proves particularly meaningful because it simplifies the control object to a large extend.[7]

3.2 Subspace identification method (SIM)

SIM incorporates system theory, linear algebra and statistics. It estimates states from input-output data using linear algebra tools (QR factorization and SVD). Once these states are known, identification becomes a linear least square problem and system matrices are easily determined. SIM provides a low-order model directly from input-output data, without having to compute the high-order model, which is useful for multivariable systems of large dimension in industry. [8] Besides, procedures of SIM are simple, and include no iterative optimization, hence, no convergence issue. The only user-specified parameter is the order of the model, which can be determined by SVD.[7,8]

4. Subspace identification method

4.1 Notation

4.1.1 System related Matrices. The extended observability matrix Γ_i and extended controllability matrix Δ_i^d are used to determine the system matrices, and are respectively defined as:

$$\Gamma_i \triangleq \begin{pmatrix} C & CA & CA^2 & \cdots & CA^{i-1} \end{pmatrix}^T \in \mathbb{R}^{li\times n} \tag{1}$$

$$\Delta_i^d \triangleq \begin{pmatrix} A^{i-1}B & A^{i-2}B & \cdots & AB & B \end{pmatrix} \in \mathbb{R}^{n\times mi} \tag{2}$$

The lower block triangular Toeplitz matrix H_i^d from the iteration of state equations, is defined as:

$$H_i^d \triangleq \begin{bmatrix} D & 0 & \cdots & 0 \\ CB & D & \cdots & 0 \\ CAB & CB & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{i-2}B & CA^{i-3}B & \cdots & D \end{bmatrix} \in \mathbb{R}^{li \times mi} \tag{3}$$

4.1.2 Block Hankel Matrices. The input block Hankel matrix is constructed by past inputs U_p and future inputs U_f . i is user-defined and should be at least larger than the maximum order of the system to be identified. j is typically equal to $s-2i+1$ for given data of length s. The output block Hankel matrix $Y_{0|2i-1} = \begin{bmatrix} Y_p \,|\, Y_f \end{bmatrix}^T$ is obtained from outputs similarly.

$$U_{0|2i-1} \triangleq \left[\begin{array}{cccc} u_0 & u_1 & \cdots & u_{j-1} \\ u_1 & u_2 & \cdots & u_j \\ \vdots & \vdots & \vdots & \vdots \\ u_{i-1} & u_i & \cdots & u_{i+j-2} \\ \hline u_i & u_{i+1} & \cdots & u_{i+j-1} \\ \vdots & \vdots & \cdots & \vdots \\ u_{2i-1} & u_{2i} & \cdots & u_{i+j-2} \end{array}\right] = \begin{bmatrix} U_{0|i-1} \\ U_{i|2i-1} \end{bmatrix} = \begin{bmatrix} U_p \\ U_f \end{bmatrix} = \left[\begin{array}{cccc} u_0 & u_1 & \cdots & u_{j-1} \\ \vdots & \vdots & \vdots & \vdots \\ u_{i-1} & u_i & \cdots & u_{i+j-2} \\ \hline u_i & u_{i+1} & \cdots & u_{i+j-1} \\ u_{i+1} & u_{i+2} & \cdots & u_{i+j} \\ \vdots & \vdots & \cdots & \vdots \\ u_{2i-1} & u_{2i} & \cdots & u_{i+j-2} \end{array}\right] = \begin{bmatrix} U_{0|i} \\ U_{i+1|2i-1} \end{bmatrix} = \begin{bmatrix} U_p^+ \\ U_f^- \end{bmatrix} \tag{4}$$

4.2 Procedures

SIM mainly involves two steps: 1) estimation of state sequence from input-output data by using algebra tools (QR, SVD); 2) least-squares estimation of state-space matrices.
Based on the iteration of state equations, we have

$$Y_{0|i-1} = \Gamma_i X_0 + H_i U_{0|i-1} \tag{5}$$

$$Y_{i|i2-1} = \Gamma_i X_i + H_i U_{i|2i-1} \tag{6}$$

From state equations, future state can be rewritten as in the row space of past inputs and past outputs:

$$X_i = A^i X_0 + \Delta_i U_{0|i-1} = A^i(-\Gamma_i^+ H_i U_{0|i-1} + \Gamma_i^+ Y_{0|i-1}) + \Delta_i U_{0|i-1} = \begin{bmatrix} \Delta_i - A^i \Gamma_i^+ H_i & A^i \Gamma_i^+ \end{bmatrix} \begin{bmatrix} U_{0|i-1} \\ Y_{0|i-1} \end{bmatrix} \triangleq L_p W_p \tag{7}$$

The oblique projection of future outputs onto past data and along the future inputs, defined as O_i , is evidently, the projection is the product of extended observability matrix and future states.

$$\mathrm{O}_i = Y_f /_{U_f} W_p = (Y_{i|2i-1} / U_{i|2i-1}^{\perp})(W_p / U_{i|2i-1}^{\perp})^{\dagger} W_p = \Gamma_i L_p W_p = \Gamma_i W_i \tag{8}$$

Compute states from SVD of weighted matrix

$$\mathrm{O}_i = USV^T = [U_1 \quad U_2] \begin{bmatrix} \Sigma_n & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} = U_1 \Sigma_n V_1^T \tag{9}$$

Compared to (9), the extended observability matrix is equal to

$$\Gamma_i = U_1 \Sigma_n^{1/2} T \tag{10}$$

The future state is equal to

$$X_i = \Gamma_i^{\dagger} \mathrm{O}_i = \Gamma_i^{\dagger} \cdot Y_i /_{U_{i|2i-1}} W_p \tag{11}$$

The system matrices A, B, C and D can be solved from (12) in a linear least-square sense.

$$\begin{bmatrix} X_{i+1} \\ Y_{i|i} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} X_i \\ U_{i|i} \end{bmatrix} \tag{12}$$

5. Identification of traction network

5.1. Design of experiment

A 10-conductor traction network structure of 30km is built on MATLAB/Simulink as the real system to identify. Segments of every 5km are represented by equivalent π-circuits. Input signals are the same for all experiments and is zero-mean Gaussian white noise sequence with variance of 1. Because studies for the traction power supply system generally consider up to the 50th harmonics of 2500 Hz, we choose 0.00004s as sampling period. The length of collected data is 10000.

5.2. Identification results

As there are families of models depending on order decision, we adopt a commonly used performance indicator, mean related variance (MVAF), defined as

$$MVAF(\%) = \frac{1}{l}\sum_{j=1}^{l}(1 - \mathrm{var}(y - \hat{y}) / \mathrm{var}(y)) \tag{13}$$

where y is the real output and $\hat{y}$ is the output estimated by the obtained model.

5.2.1 Normal operation. First we suppose the train is on the up track, 30 km from the substation and introduce the Gaussian signals at the primary-side of substation as inputs. Considering the singular value spectrum from SVD of oblique projection in figure 4, high orders represented by very small eigenvalues tend to have minor influence on the response of system, therefore, can be ignored with acceptable accuracy level. Thus we assume that a reduced model of 4-order, or even 2-order can reflect majority of the system's response to given inputs.

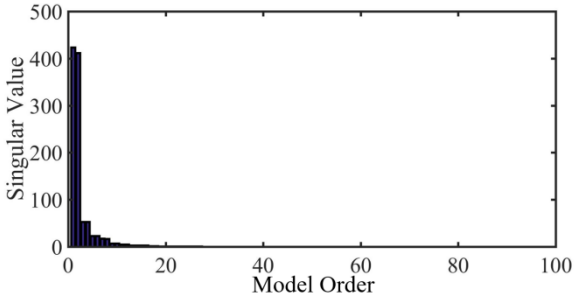


Figure 4. Obvious large gaps appear between the second and the third singular values, but from the fourth, all the rest singular value spectrum decreases more continuously. So a 4-order or 2-order model may well contain the majority of system's characteristics.

With the order of system chosen as 2, the identification result of system matrices is

$$A = \begin{bmatrix} 0.9216 & 0.3884 \\ -0.3802 & 0.9218 \end{bmatrix} B = \begin{bmatrix} -0.0145 \\ -0.0079 \end{bmatrix} C = \begin{bmatrix} -2.3957 & 1.9604 \end{bmatrix} D = 8.7228 \times 10^{-4} \tag{14}$$

With the order of system chosen as 4, the identification result of system matrices is

$$A = \begin{bmatrix} 0.9216 & 0.3884 & 0.0074 & -0.0032 \\ -0.3802 & 0.9218 & -0.0043 & -9.1749\times10^{-4} \\ -0.0010 & -0.0017 & -0.0902 & 0.9953 \\ 0.0015 & 6.3553\times10^{-4} & -0.9922 & -0.0879 \end{bmatrix}$$
$$B = \begin{bmatrix} -0.0145 & -0.0078 & -0.0057 & 0.0022 \end{bmatrix}^{\mathrm{T}}$$
$$C = \begin{bmatrix} -2.3957 & 1.9604 & 1.1139 & -0.2144 \end{bmatrix} \tag{15}$$
$$D = 0.0010$$

Figure 5 shows that outputs from the 2-order and 4-order identified systems correspond well to those of simulation model. 4-order model restores the intensive, slight variance more accurately compared to 2-order model as it captures more system's information. With the same inputs, MVAF of 2-order model is 97.8057%. Likewise for other higher orders, the accuracy is higher than 99% and increases very slightly with orders, as shown in table 1. In the frequency domain, we mainly focus on the

characteristics between 1Hz and 2500 Hz, in which train-network coupling problems appear. Figure 6 shows that 2-order and 4-order models can replicate the behaviour of real system to a large extent.

Table 1. Accuracy of identification results of different orders.

	2-order	3-order	4-order	5-order	6-order	7-order	8-order	9-order	10-order
MVAF(%)	97.8057	97.8238	99.4169	99.4097	99.6090	99.6808	99.9550	99.9549	99.9330

We then set the sinusoid voltage for real-life operating system as inputs. The two identification models can both give very close sinusoidal voltages at train's position compared to the original system. For peak values in figure 7, the bias is 1.84% for the 2-order system, and 0.26% for the 4-order system. We consider the small deviation negligible for depicting the performance of traction network in operation.

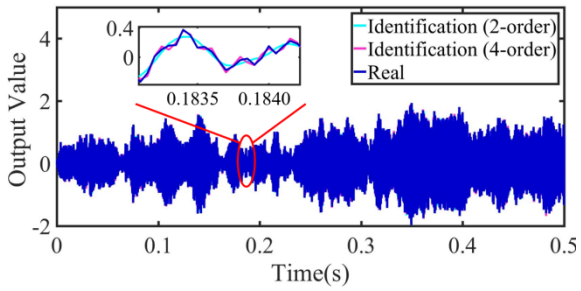


Figure 5. Outputs from simulation model and identified model (order=2 and order=4)

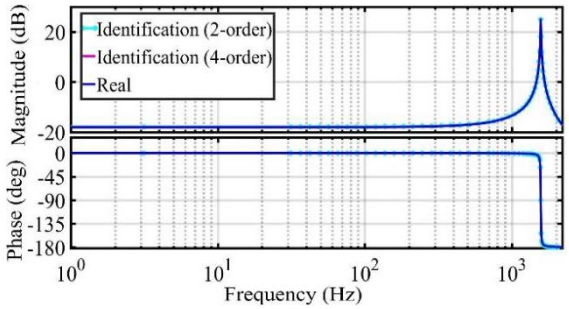


Figure 6. Bode diagram of real system and identification results

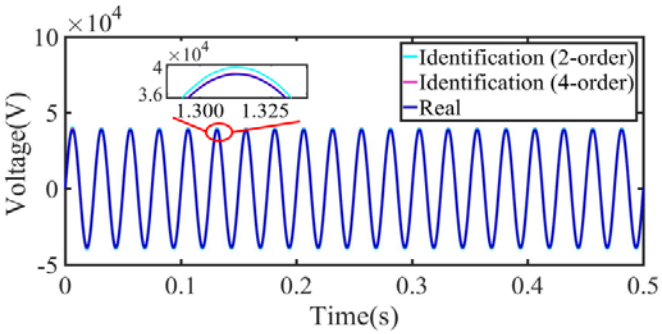


Figure 7. Using sinusoid voltages under normal operation situation as inputs, the peak value of outputs 38930V for real system is, 39950V for 2-order model and 39050V for 4-order model.

Since the high-speed train could be running at different places, the same experiments are conducted at different locations along the traction network with wire voltages at train's location as outputs. MVAF of up track and down track identification results are listed in table 2 and table 3. 4-order can guarantee MVAF higher than 99%, so considering the need of simplicity in mechanism analysis, we assume 4-order as a uniform choice for the reduced model of traction network in normal operation.

Table 2. MVAF(%) with train at different distances from substation on up track.

	0km	5km	10km	15km	20km	25km	30km
2-order	94.0741	94.5991	97.3786	99.4645	99.3250	98.3506	97.8043
4-order	99.3608	99.5146	99.4546	99.7636	99.0910	99.0820	99.4194

Table 3. MVAF(%) with train at different distances from substation on down track.

	0km	5km	10km	15km	20km	25km	30km

2-order	94.1072	94.5740	97.3714	99.4670	99.3244	98.3690	97.7996
4-order	99.3636	99.5185	99.4552	99.7652	99.0905	99.0903	99.4212

5.2.2 Fault Cases. Once wire malfunction occurs on up track, including faults like T-R short circuit, F-R short circuit and T-F shout circuit or any wire break, protective breakers will be consecutively activated to cut off all the wires on up track. Wires in down track can still work with half structure.[9] If any AT falls into faulty state, similarly, breakers will cut off its connection with feeder, traction line and rail. The action would not interfere with the rest of traction network. Considering the symmetry of traction network, we assume network models with different numbers of faulty AT. AT1 and AT2 are on up track, respectively at 15km and 30km. AT3 and AT4 are on the down track, respectively at 15km and 30km. We construct new reduced mathematical models after protection actions. Table 4 shows that 4-order identification results all maintain high accuracy to reflect the performances of real systems in different fault cases.

Table 4. MVAF(%) of identification results under different fault cases.

	AT	AT1, AT3	AT2, AT4	AT1, AT2, AT3, AT4	Wire Faults
2-order	97.2415	96.8815	97.4644	96.6052	97.6595
4-order	99.6704	99.7713	99.4829	99.0705	97.8715

With both 4-order mathematical models in normal and fault cases, the variation of network states can be seen as parameter changes of 4-order system matrices. Such unification can be applied to detect the fluctuation range of traction network, and to study its influence on coupling mechanism.

6. Conclusion

In order to solve limitation of the existing modelling methods of traction network in mechanism analysis, this paper applied subspace identification to construct reduced mathematical models. Experiments showed that with acceptable accuracy, a uniform order decision was applicable for the identification of large-scale traction network. The variation of network states, including location variation and faults, could be seen as the parameter deviation of 4-order system matrices without changing orders, which largely facilitates further analyses. (1) Once the low-order mathematical models are obtained, they can be used to predict or analyze the dynamic performance for given certain abnormal inputs, such like perturbation of primary side of traction substation. (2) The simple model of traction network can be combined with the train's model to conduct more compatible coupling stability criteria under multiple operation states of traction network. (3) In accordance with the idea of robust control, the variances of system matrices for these 4-order models can be concluded as uncertainty of a nominal model. This structure would improve the robustness of corresponding control designs, especially when dealing with fault cases.

7. References

[1] Hu H, Tao H, Wang X, Blaabjerg F, He Z and Gao S 2018 *IEEE Trans. on Power Electro.* **33** 4627-42
[2] Wang H, Wu M and Sun J 2015, *IEEE Trans. on Power Electron.* **30** 5318-5330
[3] Liao Y, Liu Z, Zhang H, and Wen B 2018 *IEEE Trans. on Ind. Appli.* **54** 4999-5011
[4] Lü X and Wang X 2017 *Proc. of the CSEE* **3** 217-229
[5] Song W, Feng X, Ge X, and Cui H 2016 *IET Electrical Systems in Transportation* **6** 88-95
[6] Van Overschee P and De Moor B 1996 *Subspace Identification for Linear Systems* (US: Springer)
[7] Katayama T 2005 *Subspace Method for System Identification* (US: Springer)
[8] Zhang G, Liu Z, Yao S, Liao Y and Xiang C 2016 *IEEE Trans. on Transpor. Elec.* **2** 244-255

[9] Tan X 2015 *Protective Relay of Electrical and Traction Power Supply Systems for High-speed Railways* (Chengdu: Southwest Jiaotong University Press)

Acknowledgments
This work was supported by the NSFC (No.61733015). A deep gratitude for our colleagues from Institute of Systems Science and Technology at Southwest Jiaotong University who provided insight and expertise that greatly assisted the research.