

Supplementary Files

S1. Convergence and computational complexity

This section presents the convergence and time complexity of NAE. The convergence of this algorithm is determined by calculating the loss value of objective function in each iteration of the model. Taking the top 100 networks in the real IAV dataset as an example, Fig. S1 shows the convergence curves of solving the mathematical programming responding to the 100 networks individually, where the x -axis and y -axis represent the number of iterations in each network and the value of loss function, respectively. Therefore, most curves can quickly converge to a small loss value, and the maximum number of iterations does not exceed 50 to complete the convergence task. (13), (17-19), (20), and (23) need to be updated during iterative optimization, and their computational complexities are $O(\max(n^2, mn))$, $O(n^2)$, $O(n^2)$, and $O(n^2)$, respectively. Thus, the total computational complexity of NAE is $O(\max(n^2, mn))$.

In numerical experiments, this subsection analyzes the actual running time of the NAE model through DREAM3 and IAV case studies. The experimental results are shown in the following Fig. S2. Fig. S2(a) shows the running time of each data in the DREAM 3 dataset for 1000 times. The size of the gene network for each evaluation is 10, and the evaluation time of most networks is within 0.01s-0.02s. Fig. S2(b) shows the running time of 184 pathways in the IAV dataset (averaged over 1000 runs), the number of genes in each pathway is shown on the left, and the histogram is the running time of the corresponding algorithm.

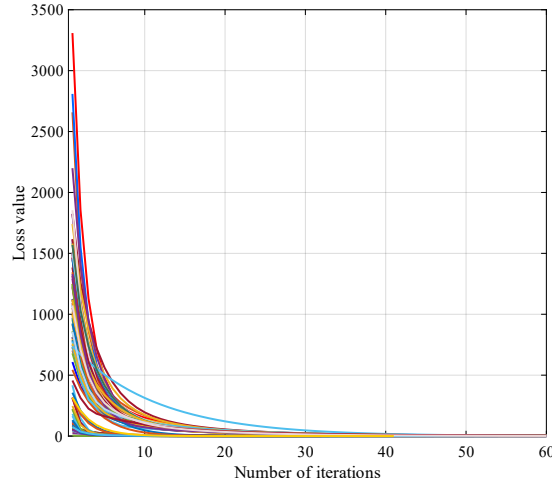


Fig. S1. Convergence curves of the top 100 networks in the IAV dataset.

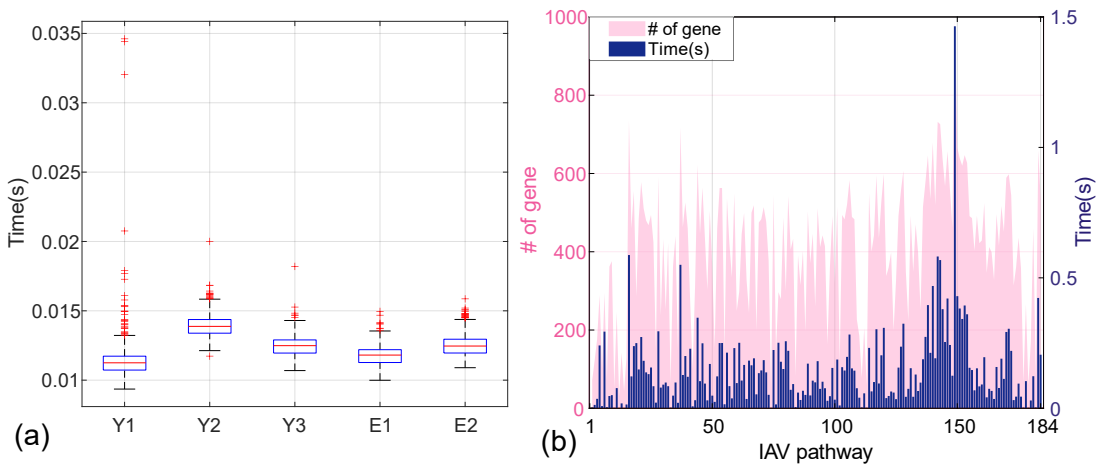


Fig. S2. Running time analysis. (a) DREAM 3 datasets; (b) IAV datasets.

S2. Evaluating NAE model on the DREAM4 datasets

The evaluation results of the NAE model on the DREAM 4 datasets are listed in Table S1, and the results are consistent with those in DREAM 3 datasets.

TABLE S1.
EXPERIMENTAL RESULTS ON DREAM4 DATASETS

	10_1	10_2	10_3	10_4	10_5	Mean
$p = 1$	0.088	0.001	0.023	0.043	0.059	0.043
$p = 2$	0.063	0.210	0.027	0.076	0.214	0.118
$p = \frac{1}{3}$	0.009	0.028	0.033	0.017	0.173	0.052
$p = \frac{1}{2}$	0.150	0.236	0.028	0.059	0.045	0.104
$p = \frac{2}{3}$	0.241	0.049	0.005	0.044	0.032	0.074
$p = 0$	0.008	0.126	0.004	0.014	0.007	0.032

S3. Analysis of key modules

The NAE model uses DBN to simulate the dynamic gene responses, and calculates the activity values of different networks. Furthermore, prior-knowledge constraints and regularization penalties are added to strengthen the model's ability to link regulators and genes. Next, we analyzed the contribution of the regularization penalty term, the constraint term, and the underlying DBN objective function in NAE through numerical experiments. The results are listed in the following Table I. As shown in Table S2, the experiments show that both the constraint term and the regularization penalty term have a positive impact on the proposed NAE model. Relatively, the regularization penalty term has a slightly larger impact. In addition, removing the regularization penalty term and the constraint term at the same time significantly reduces the model ability of evaluating network activity.

TABLE S2.
ANALYSIS OF KEY MODULES THROUGH THE DREAM3 DATASETS

	Y1	Y2	Y3	E1	E2	Mean
NAE	0.020	0.002	0.029	0.020	0.001	0.014
NAE (No regularization)	0.017	0.442	0.046	0.387	0.001	0.179
NAE (Unconstrained)	0.008	0.081	0.218	0.174	0.001	0.096
NAE (DBN emulation only)	1.000	0.001	0.001	1.000	0.001	0.4006

S4. Significant pathways

Tables S3-S8 list the significant KEGG pathways (gene sets) identified by the comparing six methods (GSA does not enrich any significant pathway). For simplicity, we show here with a maximum of 5 pathways per method. 9, 4, 22, 3, 3 and 18 significant pathways are identified by the six methods, respectively.

TABLE S3.
SIGNIFICANT KEGG PATHWAYS AND DESCRIPTIONS IN IAV VIA NAE+miRNA.

Pathway	FDR p-value	Name	Description
hsa03013	0.023	Nucleocytoplasmic transport	The exchange of molecules between the nucleus and cytoplasm is mediated through nuclear pore complexes (NPCs) embedded in the nuclear envelope.
hsa03030	0.023	DNA replication	A complex network of interacting proteins and enzymes is required for DNA replication.
hsa03440	0.023	Homologous recombination	Homologous recombination (HR) is essential for the accurate repair of DNA double-strand breaks (DSBs), potentially lethal lesions.
hsa03460	0.023	Fanconi anemia pathway	The Fanconi anemia pathway is required for the efficient repair of damaged DNA, especially interstrand cross-links (ICLs).
hsa04120	0.023	Ubiquitin mediated proteolysis	Protein ubiquitination plays an important role in eukaryotic cellular processes.

Note: Bold represents IAV-related pathways. Only the top 5 of the 9 significant pathways are listed.

TABLE S4.
SIGNIFICANT KEGG PATHWAYS AND DESCRIPTIONS IN IAV VIA NAE.

Pathway	FDR p-value	Name	Description
hsa04110	0.046	Cell cycle	Cyclin-dependent kinases (CDKs) are key regulatory enzymes, each consisting of a catalytic CDK subunit and an activating cyclin subunit.
hsa05163	0.046	Human cytomegalovirus infection	Human cytomegalovirus (HCMV) is an enveloped, double-stranded DNA virus that is a member of beta-herpesvirus family.
hsa05225	0.046	Hepatocellular carcinoma	Hepatocellular carcinoma (HCC) is a major type of primary liver cancer
hsa05323	0.046	Rheumatoid arthritis	Rheumatoid arthritis (RA) is a chronic autoimmune joint disease where persistent inflammation affects bone remodeling leading to progressive bone destruction.

TABLE S5.
SIGNIFICANT KEGG PATHWAYS AND DESCRIPTIONS IN IAV VIA NETGO.

Pathway	FDR p-value	Name	Description
hsa03013	0.010	Nucleocytoplasmic transport	The exchange of molecules between the nucleus and cytoplasm is mediated through nuclear pore complexes (NPCs) embedded in the nuclear envelope.
hsa03015	0.010	mRNA surveillance pathway	The mRNA surveillance pathway is a quality control mechanism that detects and degrades abnormal mRNAs.
hsa03018	0.010	RNA degradation	The correct processing, quality control and turnover of cellular RNA molecules are critical to many aspects in the expression of genetic information.
hsa03020	0.010	RNA polymerase	Genetic information processing.
hsa03022	0.010	Basal transcription factors	Genetic information processing.

Note: Only the top 5 of the 22 significant pathways are listed.

TABLE S6.
SIGNIFICANT KEGG PATHWAYS AND DESCRIPTIONS IN IAV VIA SIGPATHWAY.

Pathway	FDR p-value	Name	Description
hsa04110	0.001	Cell cycle	Cyclin-dependent kinases (CDKs) are key regulatory enzymes, each consisting of a catalytic CDK subunit and an activating cyclin subunit.
hsa03040	0.001	Spliceosome	After transcription, eukaryotic mRNA precursors contain protein-coding exons and noncoding introns.
hsa03030	0.001	DNA replication	A complex network of interacting proteins and enzymes is required for DNA replication.
hsa04141	0.001	Protein processing in endoplasmic reticulum	The endoplasmic reticulum (ER) is a subcellular organelle where proteins are folded with the help of luminal chaperones.
hsa04137	0.002	Mitophagy	Mitochondria act as the energy powerhouse of the cell, and are essential for eukaryotic cells to grow and function normally.

TABLE S7.
SIGNIFICANT KEGG PATHWAYS AND DESCRIPTIONS IN IAV VIA GSEA.

Pathway	FDR p-value	Name	Description
hsa03020	0.018	RNA polymerase	Genetic information processing.
hsa03040	0.022	Spliceosome	After transcription, eukaryotic mRNA precursors contain protein-coding exons and noncoding introns.
hsa04115	0.035	p53 signaling pathway	p53 activation is induced by a number of stress signals, including DNA damage, oxidative stress and activated oncogenes.

TABLE S8.
SIGNIFICANT KEGG PATHWAYS AND DESCRIPTIONS IN IAV VIA GSVA.

Pathway	FDR p-value	Name	Description
hsa03018	0.038	RNA degradation	The correct processing, quality control and turnover of cellular RNA molecules are critical to many aspects in the expression of genetic information.
hsa05120	0.038	Epithelial cell signaling in Helicobacter pylori infection	Two major virulence factors of H. pylori are the vacuolating cytotoxin (VacA) and the cag type-IV secretion system (T4SS) and its translocated effector protein, cytotoxin-associated antigen A (CagA).
hsa05146	0.038	Amoebiasis	Entamoeba histolytica, an extracellular protozoan parasite is a human pathogen that invades the intestinal epithelium.