A. Analog components, modules, and necessary ports for graph generation

To provide a comprehensive overview, the following section details how we transform circuits into graph data structures for processing. This method involves dividing the circuit into multiple simple modules (such as differential pairs, current mirror loads, etc.) and treating these modules as nodes in the graph. Each node contains both discrete and continuous parts.

A.1. One-Hot Encoding for Device Types

The discrete part of the node matrix $\boldsymbol{V}$ is formed using one-hot encoding based on the device type.

In this study, we adopted a one-hot encoding method to represent different types of devices in the circuit. This encoding allows us to denote each device as a unique one-hot vector in a high-dimensional space. The following table provides the complete one-hot encodings for all specified device types.

Table 1. Complete one-hot encodings for all specified device types.

Device Type	One-Hot Encoding
NMOS (3 terminals)	[1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]
PMOS (3 terminals)	[0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]
Current_mirror_PMOS (3 terminals)	[0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]
Current_mirror_NMOS (3 terminals)	[0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]
diff_pair_NMOS (5 terminals)	[0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]
diff_pair_PMOS (5 terminals)	[0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]
bias_pair_NMOS (5 terminals)	[0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]
bias_pair_PMOS (5 terminals)	[0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]
Resistor (2 terminals)	[0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]
Capacitor (2 terminals)	[0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0]
Vdd	[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0]
Vss	[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0]
Vout	[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0]
VINN	[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0]
VINP	[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0]
IBIASP	[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0]
IBIASN	[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0]
IBIASP_2	[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0]
IBIASN_2	[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1]

Each one-hot encoding represents a unique device type within the circuit, allowing for clear and unambiguous identification in the graph representation.

A.2. Port Information for Different Device Types

To provide a comprehensive overview, the following table details the port information for various device types. This information is critical for correctly modeling the connections within the circuit graph.

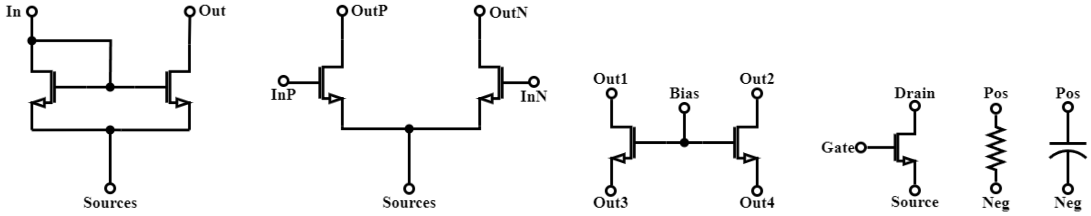
The port sequences listed in Table 2 are used to define the connections between devices in the circuit. For example, in a MOSFET (either NMOS or PMOS), the sequence [Drain, Gate, Source] specifies that the first element in any connection matrix corresponds to the drain terminal, the second to the gate, and the third to the source.

The images of these devices and their ports are drawn in the Figure1.

The edges between nodes represent the connections between different modules. The attributes of each edge are represented by a matrix $\boldsymbol{\xi} \in \mathbb{R}^{k \times k}$, where k is the maximum port number of all types of components. When $\boldsymbol{\xi}_{i,j} = 1$, it indicates that the i-th port of the first device is connected to the j-th port of the second device. Conversely, if $\boldsymbol{\xi}_{i,j} = 0$, it indicates no connection between these ports.

Table 2. Port sequences for different device types.

Device Type	Port Sequence
MOSFETs [NMOS, PMOS]	[Drain, Gate, Source]
Current_Mirror_[NMOS, PMOS]	[In, Out, Sources]
Differential_Pair_[NMOS, PMOS]	[InP, InN, OutP, OutN, Sources]
Bias_Pair_[NMOS, PMOS]	[Bias, Out1, Out2, Out3, Out4]
Resistor	[Pos, Neg]
Capacitor	[Pos, Neg]
Vdd	[Vdd]
Vss	[Vss]
Vout	[Vout]
VINN	[VINN]
VINP	[VINP]
IBIASP	[IBIASP]
IBIASN	[IBIASN]
IBIASP_2	[IBIASP_2]
IBIASN_2	[IBIASN_2]



(a) (b) (c) (d) (e) (f)

Figure 1. Device Type Example.(a)Current_Mirror_NMOS.(b)Differential_Pair_NMOS.(c)Bias_Pair_NMOS.(d)NMOS.(e)Resistor.(f)Capacitor.

The continuous part of $\boldsymbol{V}$ includes specific parameter values for each module. For a differential pair, the continuous part is the width and length (W, L) of the MOS transistors. For a compensation capacitor, the continuous part is its capacitance value.

A.3. Normalization and Sample Information

In this work, the design of our operational amplifiers took into account a comprehensive set of 13 design metrics. Table 3 provides the normalized parameters for these design metrics.

To ensure a consistent number of sampling points for the device parameters at each order of magnitude, our sampling parameters are as follows: w(The MOSFET channel width) : $[1e-6, 2e-6, 5e-6, 1e-5, 2e-5, 5e-5, 1e-4]$, r(The resistor value) : $[1e4, 2e4, 5e4, 1e5, 2e5, 5e5, 1e6]$, and c(The capacitor value) : $[1e-12, 2e-12, 5e-12, 10e-12]$. Furthermore, to ensure that the distribution of the sampled data is within $[0, 1]$ and uniformly distributed within this interval, we perform a linear transformation after taking the logarithm of the device parameters and finally map them into $[0, 1]$. Taking MOSFET as an example, its sampling space is $[1e-6, 1e-4]$. After taking the logarithm, it is $[-6, -4]$. Then, by dividing by 2 and adding 3, the interval becomes $[0, 1]$.

Table 3. Performance Metrics and Their Normalization Parameters

Metric	Normalization Parameter
Power Consumption(P_{diss})	1×10^{-3}
DC Gain($Gain_{DC}$)	100
Gain-Bandwidth Product(GBW)	10×10^{6}
Phase Margin(PM)	180
Positive Slew Rate(SR_P)	10×10^{6}
Negative Slew Rate(SR_N)	10×10^{6}
Output Voltage Swing Low(VOL)	1.2
Output Voltage Swing High(VOH)	1.2
Common-Mode Rejection Ratio($CMRR$)	100
Power Supply Rejection Ratio($PSRR$)	100
Input Equivalent Noise @ 1 kHz($Noise_{@1kHz}$)	1×10^{-6}
Input Equivalent Noise @ 1 GHz($Noise_{@1GHz}$)	1×10^{-7}
Load Capacitance(C_L)	10×10^{-12}

A.4. Example

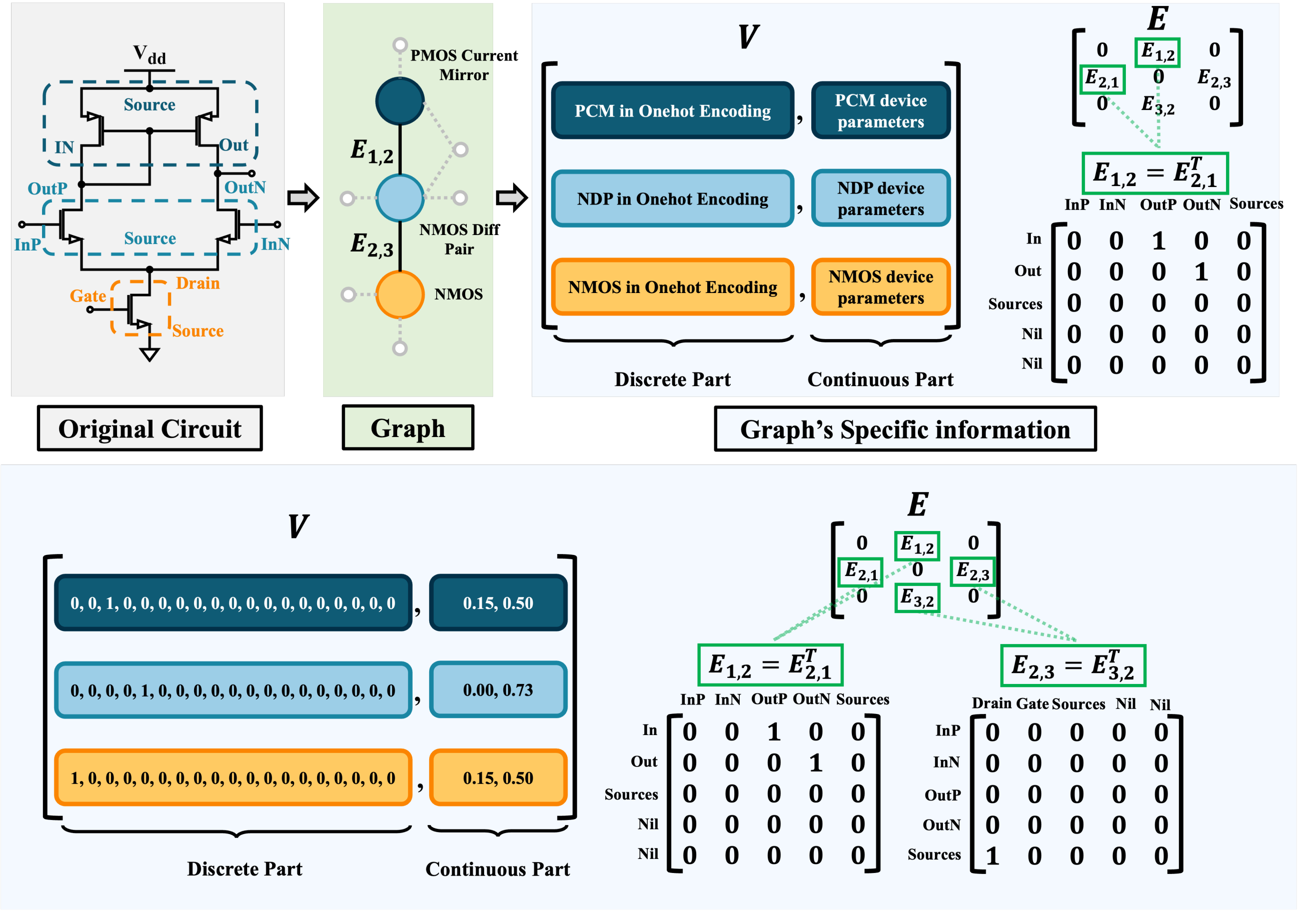


Figure 2. A graph example.

Consider the NMOS-input 5-transistor Operational Transconductance Amplifier (OTA) mentioned in Figure2. The circuit can be divided into:

- A PMOS Current Mirror (PCM) with one-hot encoding [0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]
- An NMOS Differential Pair (NDP) with one-hot encoding [0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]
- A single NMOS transistor with one-hot encoding [1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]

Each component is one-hot encoded to form the discrete part of the node matrix $\boldsymbol{V}$.

In the context of a five-transistor OTA (5-T-OTA) example, the continuous part of the node matrix $\boldsymbol{V}$ includes specific parameters such as the channel width (W) and length (L) of MOS transistors. These parameters are crucial for defining the operational characteristics of each module within the circuit.

According to the logarithmic normalization method, the continuous part of $\boldsymbol{V}$ can be described as follows:

- **PMOS Current Mirror(PCM):** The PMOS transistors in the current mirror might have W/L ratios of $10\mu m/2\mu m$. After applying the normalization factor, this translates to $W = 3 + 1/2 * log_{10}(10 * 10^{-6}) = 0.50$, $L = 3 + 1/2 * log_{10}(2 * 10^{-6}) = 0.15$.
- **NMOS Differential Pair(NDP):** The NMOS transistors in the differential pair might have W/L ratios of $30\mu m/1\mu m$. After applying the normalization factor, this translates to $W = 3 + 1/2 * log_{10}(30 * 10^{-6}) = 0.73$, $L = 3 + 1/2 * log_{10}(1 * 10^{-6}) = 0$.
- **Single NMOS Transistor:** The single NMOS transistor might have a W/L ratio of $10\mu m/2\mu m$. After applying the normalization factor, this translates to $W = 3+1/2*log_{10}(10*10^{-6}) = 0.50$, $L = 3+1/2*log_{10}(2*10^{-6}) = 0.15$.

The edge attributes between adjacent nodes are recorded using matrices. For unconnected ports, the element corresponding to the edge matrix is zero.

Specifically in this example,the edge attribute matrix E is a $3 \times 3 \times 5 \times 5$ tensor, where each slice represents the connection between two nodes. Specifically, we have:

$$E = \begin{bmatrix} 0 & E_{1,2} & 0 \\ E_{2,1} & 0 & E_{2,3} \\ 0 & E_{3,2} & 0 \end{bmatrix}$$

Here, $E_{1,2}$ and $E_{2,1}$ represent the connections between the PMOS Current Mirror (Node 1) and the NMOS Differential Pair (Node 2), and $E_{2,3}$ and $E_{3,2}$ represent the connections between the NMOS Differential Pair (Node 2) and the Single NMOS Transistor (Node 3).

The detailed representation of the matrices $E_{1,2}$ and $E_{2,3}$ is as follows:

$$E_{1,2} = E_{2,1}^T = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

This matrix captures the connections from the PMOS Current Mirror to the NMOS Differential Pair.The In of the PMOS Current Mirror connects to the OutP port of the NMOS Differential Pair ($\boldsymbol{\xi}_{1,3} = 1$).The Out port of the PMOS Current Mirror connects to the OutN port of the NMOS Differential Pair ($\boldsymbol{\xi}_{2,4} = 1$).

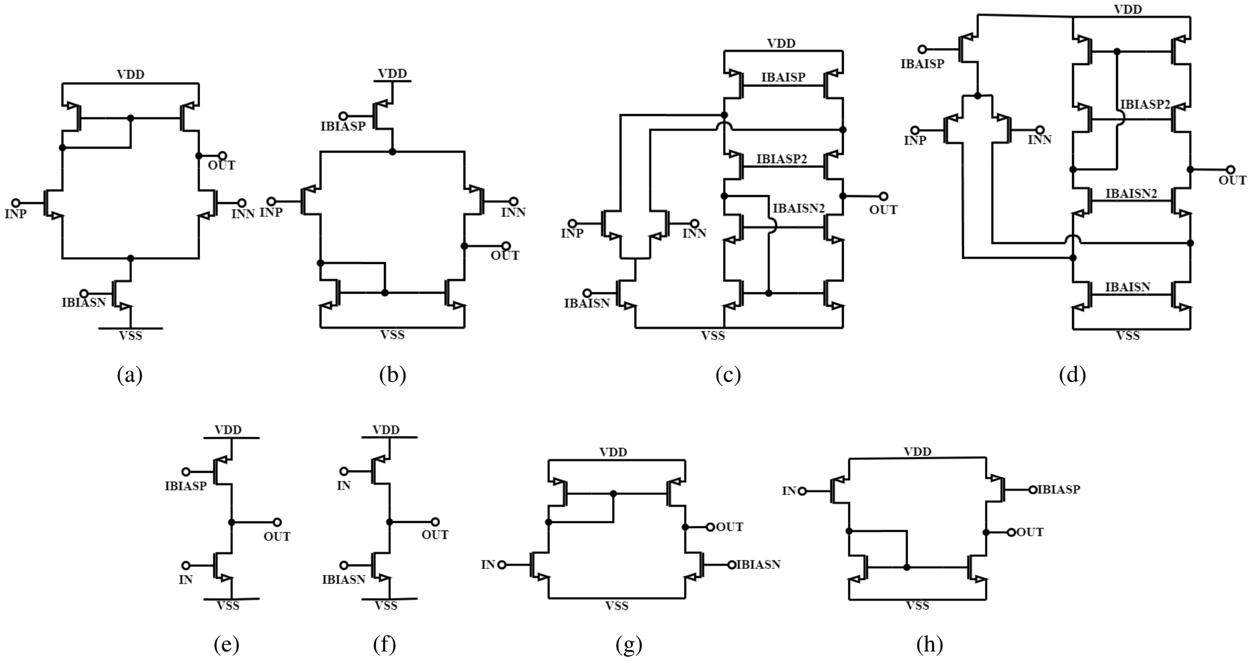
$$E_{2,3} = E_{3,2}^T = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

This matrix captures the connections from the NMOS Differential Pair to the Single NMOS Transistor.The Sources port of the NMOS Differential Pair connects to the Drain port of the Single NMOS Transistor ($\boldsymbol{\xi}_{5,1} = 1$). The edge attribute matrices $E_{1,2}$, $E_{2,1}$, $E_{2,3}$, and $E_{3,2}$ capture the specific connections between the nodes in the circuit. These matrices are crucial for accurately representing the connectivity and facilitating the simulation and analysis of the circuit.

B. Operational Amplifier

B.1. Operational Amplifier Types

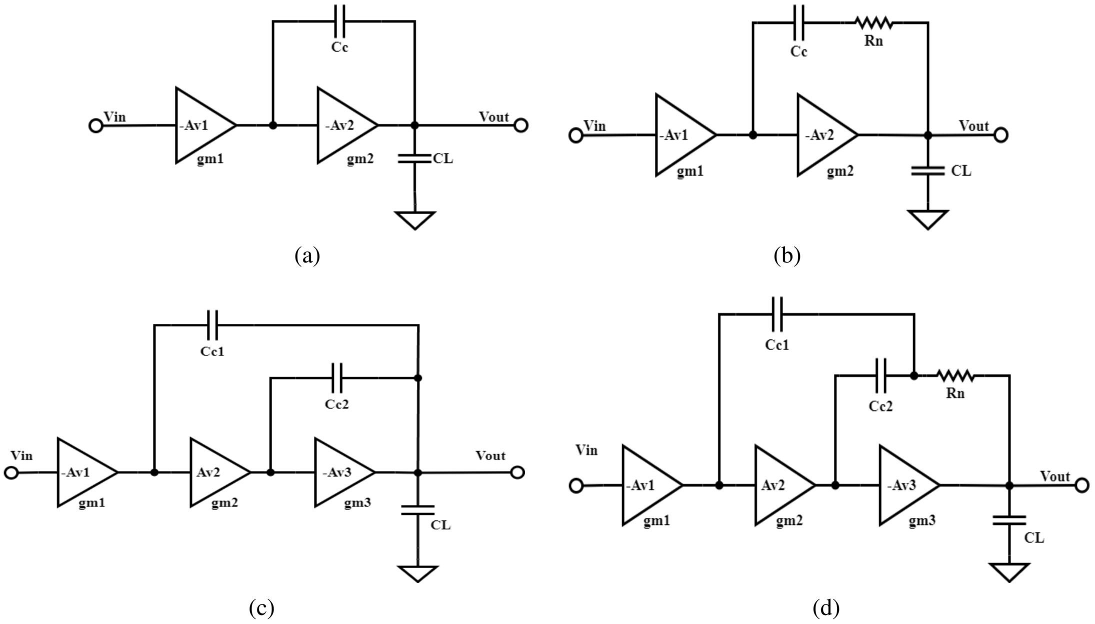
In this work, we introduce a novel data construction methodology that systematically explores the design space of analog amplifiers through the random permutation and combination of eight distinct single-stage amplifier topologies. These topologies are subsequently integrated into a comprehensive framework encompassing one single-stage amplifier structure and four multi-stage amplifier structures, namely 5-T-OTA (N&P), Folded Cascode (N&P), Common Source (N&P), and Common Source Non-Inverting (N&P). This approach yields a diverse set of transistor configuration combinations, resulting in amplifiers with varied topological and performance characteristics. Figure 3 illustrates the structures of the eight single-stage amplifiers employed in our methodology.



(a) (b) (c) (d) (e) (f) (g) (h)

Figure 3. Single-Stage-Op-amp.(a)5-T-OTA-NMOS.(b)5-T-OTA-PMOS.(c)Folded-cascode-NMOS.(d)Folded-Cascode-PMOS.(e)Common-Source-NMOS.(f)Common-Source-PMOS.(g)Non-Inverting-NMOS.(h)Non-Inverting-PMOS.

The overall operational amplifier framework encompasses Single-Stage designs, SMC (simple Miller compensation), SMCNR (Simple Miller Compensation with Nested Resistor), NMC (Nested Miller compensation), NMCNR (Nested Miller compensation with Nested Resistor). These frameworks represent various design strategies to meet specific application requirements such as high-frequency response, low power consumption, or high gain.The structures of the four types of multi-stage operational amplifiers mentioned above are shown in the Figure4.



(a)

(b)

(c)

(d)

Figure 4. Op-amp Framework.(a)SMC.(b)SMCNR.(c)NMC.(d)NMCNR.

C. Network Structure

C.1. Node prediction network

In the diffusion model, throughout the process of adding and removing noise from the graph, the number of nodes remains invariant. Therefore, before employing a denoising network, a four-layer multilayer perceptron (MLP) has to be utilized first to predict the number of nodes that satisfy the requisite structural performance. Subsequently, based on the predicted number of nodes, the denoising network is applied to the noisy graph to obtain various graphs that meet the required structural performance.

C.2. Denoising network

At the input terminal, we utilized a layer of the MLP network to map the node feature X, edge feature E, and the required structural performance y to a high-dimensional space. In the middle part, a multi-layer Graph Attention Network with variable layers was employed to update the node feature X, edge feature E, and the required structure performance y. At the output terminal, we once again applied a layer of the MLP network to map the node feature X and edge feature E back to their original dimensions.

The structure of our denoising network is illustrated in Fig.5. This network takes as input a noisy graph with node feature X and edge feature E . Simultaneously, it is guided by the normalized required structural performance y for the update process. Through this mechanism, the denoising network can predict the distribution of node feature X and edge feature E on a clean graph that adheres to the required structural performance.

C.3. Graph Attention Network

The structure of our graph attention network is depicted in Fig.6. Taking node features X, edge features E, and the normalized required structural performance y as inputs, the network updates the representations of these input node and edge features under the guidance of the normalized required structural performance y. To enhance precision and denoising capabilities, multiple Graph Attention Networks can be employed for each update.

We referred to the FiLM and PNA in Digress. In this structure, $FiLM(M1, M2) = M1{\cdot}W1 + (M1{\odot}W2){\cdot}M2 + M2$ for learnable weight matrices W1 and W2, and $PNA(X) = cat(max(X), min(X), mean(X), std(X))W$.

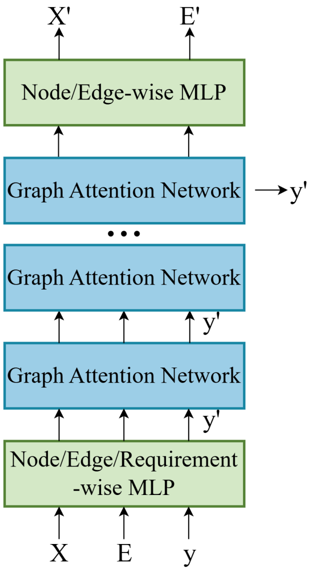


Figure 5. The structure of denoising network

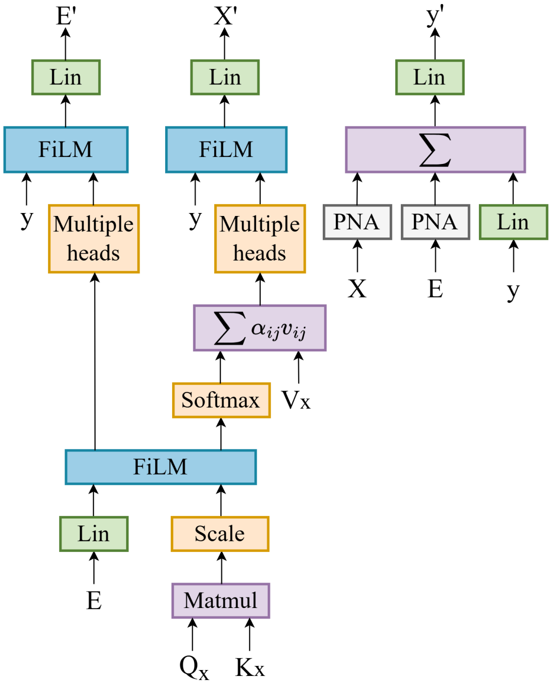


Figure 6. The structure of our graph attention network

D. Performance Evaluation

The performance metrics are divided into two categories: Category A and Category B. Category A metrics, which are preferable when larger, include $Gain_{DC}$, GBW, PM, SR_P, SR_N, VOH, $CMRR$, and $PSRR$. Category B metrics, which are preferable when smaller, include P_{diss}, VOL, $Noise@1kHz$, and $Noise_{@1GHz}$. Worth to mention that, C_L is set as the load when the amplifier is operating, which is used to examine various metrics of the circuit under this load.

Algorithm 1 Performance Evaluation

```
Input: Required metrics $Y$, Actual metrics $Y_A$, Tolerance $tol$
Initialize $Fitness = 1$
for each metric $Y_i$ in $Y$ do
    if $Y_i \in Y_{A-class}$ (A-class metrics) then
        $Fitness -= \max\left(0, (Y_i - Y_{A,i})/Y_i - tol\right) / NumOfPerformance$
    else if $Y_i \in Y_{B-class}$ (B-class metrics) then
        $Fitness -= \max\left(0, (Y_{A,i} - Y_i)/Y_{A,i} - tol\right) / NumOfPerformance$
    end if
end for
Output: $Fitness$
```

To obtain the relative error between the generated circuit metrics and the expected metrics, Algorithm 1 is applied. If a circuit metric achieves a relatively high score in this algorithm, it indicates that it better meets the required metrics.