Reinforcement Learning-based Optimal On-board Decoupling Capacitor Design Method

Hyunwook Park, Junyong Park, Subin Kim, Daehwan Lho, Shinyoung Park, Gapyeol Park, Kyungjun Cho and Joungho Kim
Terabyte Interconnection and Package Laboratory
Korea Advanced Institute of Science and Technology (KAIST)
Daejeon, Republic of Korea
hyunwookpark@kaist.ac.kr

***Abstract*— In this paper, for the first time, we propose a reinforcement learning-based optimal on-board decoupling capacitor (decap) design method. The proposed method can provide optimal decap designs for a given on-board power distribution network (PDN). An optimal decap design refers to the optimized combination of decaps at proper positions to satisfy a required target impedance. Moreover, a minimum number of decaps should be assigned for optimal decap designs. The proposed method is applied to the test on-board PDN and successfully provided 37 optimal decap designs with 4 decaps assigned each. Self impedance of PDN with the provided design satisfied the required target impedance while minimizing the number of assigned decaps.**

Keywords—Decoupling capacitor, on-board PDN, optimal design, Q-learning, reinforcement learning.

I. INTRODUCTION

Due to recent trends of the miniaturization of the electrical device, the layout design of the components on the printed circuit board (PCB) has become tight. Among the various components of the board, decoupling capacitor (decap) is one of the most important components for stable voltage supply to the integrated circuits (ICs). However, overdesigned decap designs result in additional costs, and underdesigned decap designs cause logical failures in ICs. Therefore, the optimal decap design which meets the target impedance (Z_{target}) with the minimum number of decaps is needed as shown in Fig. 1.

Various methods to optimize the on-board decap design have been studied. Decap selection algorithm based on maximum anti-resonance points and quality factors of the capacitor was proposed [1]. However, the positions of decaps are not considered and the Z_{target} is a long-standing standard. While a simple and fast method of on-board decap selection and placement was suggested [2], the process of optimization was performed by a commercial optimization tool.

In this paper, we propose a reinforcement learning (RL)-based optimal on-board decap design method. RL is the process by which an agent learns the optimal sequence of actions through the interaction with an environment. Among the RL algorithms, a Q-learning algorithm was used for the proposed method. The proposed method can design the decaps for the on-board power distribution network (PDN) based on the learned optimal policy, i.e. optimal sequence of actions. As a result, the proposed method provided optimized combinations of decaps at the proper positions to satisfy the required Z_{target} with a minimum number of decaps. The proposed method is applied to the test on-board PDN for verification and successfully provided optimal decap designs. One of the optimal decap designs is presented and compared to one of the overdesigned results.

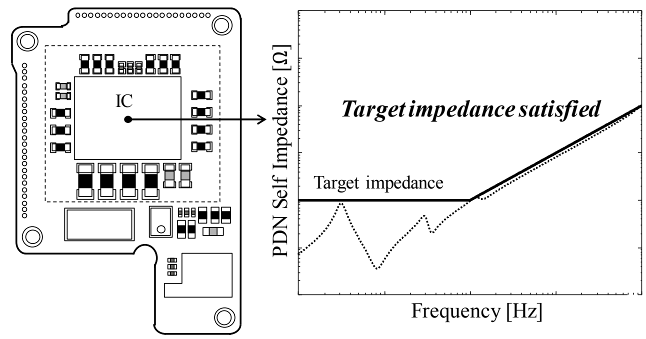

Fig. 1. Layout of PCB with the optimal decoupling capacitor design. The PDN self impedance at IC satisfies the required target impedance.

II. PROPOSED OPTIMAL ON-BOARD DECAP DESIGN METHOD

A. Concept of the proposed method

The proposed method is implemented based on Q-learning. A Q-learning algorithm can solve the problem expressed by markov decision process (MDP) consisted of states, actions, state transition probabilities, rewards and a discount factor [3]. In Q-learning, an agent defined in an environment recognizes the current state and learns how to select the action at each state that maximizes a cumulative reward received in forward time steps, i.e. Q-value [3], and the goal of the Q-learning algorithm is to learn an optimal Q-value q_{π}^{*}. Therefore, based on the optimal Q-value, the agent can obtain the optimal policy $\pi^{*}(a|s)$ to achieve one of the goal states.

Fig. 2 shows the overall concept of the proposed method. The problem to be solved is designing decaps on the on-board PDN to meet the Z_{target} with the minimum number of decaps. An agent which designs the decaps, namely decap design agent, observes the current state S_t, which means estimating the PDN

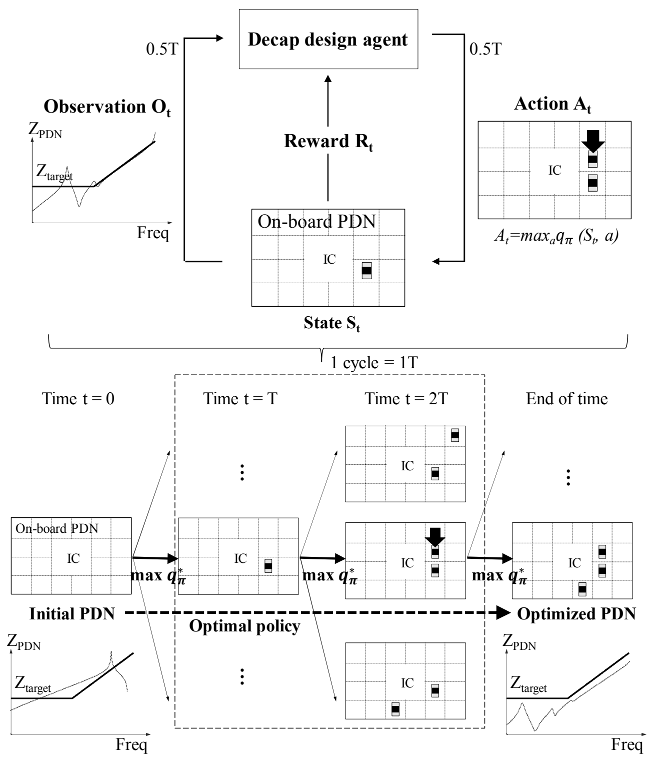

Fig. 2. Concept of the proposed Q-learning-based optimal decap design method.

impedance from the current PDN structure. Based on whether the estimated PDN impedance satisfies the Z_{target}, the agent obtains a reward R_t. Then, the agent assigns a certain type of decap on the proper position, which is denoted as an action A_t, to maximize the Q-value. For this decap design problem, Q-value is a numerical value representing the degree of satisfying the Z_{target} with the minimum number of decaps in the forward time steps. By using the Q-learning algorithm, the decap design agent can learn the optimal Q-value. Based on the optimal Q-value, the agent finally obtains the optimal policy which is the decap design guideline from the initial PDN to the decap design optimized PDN.

In Fig. 3, details about the definitions of the parameters in decap design problem are represented. The state can be defined as the PDN structure which is represented by a unit cell array and a decap array in matrix forms as shown in Fig. 3 (a). The unit cell array shows the arrangement of unit cells configuring the PDN where 0, 1, and 2 represent the absence of unit cell, a unit cell with a port, and a unit cell with a decap respectively. The decap array is the matrix that represents what type of decap is assigned on the PDN. An observation O_t is the estimation of the PDN self impedance (Z_{11}) from the current state S_t by using a segmentation method. Segmentation method is a fast and less computational method to combine each of the PDNs together [4]. Fig. 3 (b) shows the definition of the action which is assigning a certain type of decap at a certain position. Assigning a decap is also performed by the segmentation method. The reward is assigned to the agent based on the comparison between the PDN self impedance at IC and the Z_{target}. Reward R_t equals 1 if the Z_{target} is satisfied, and it is 0 if not as shown in Fig. 3 (c).

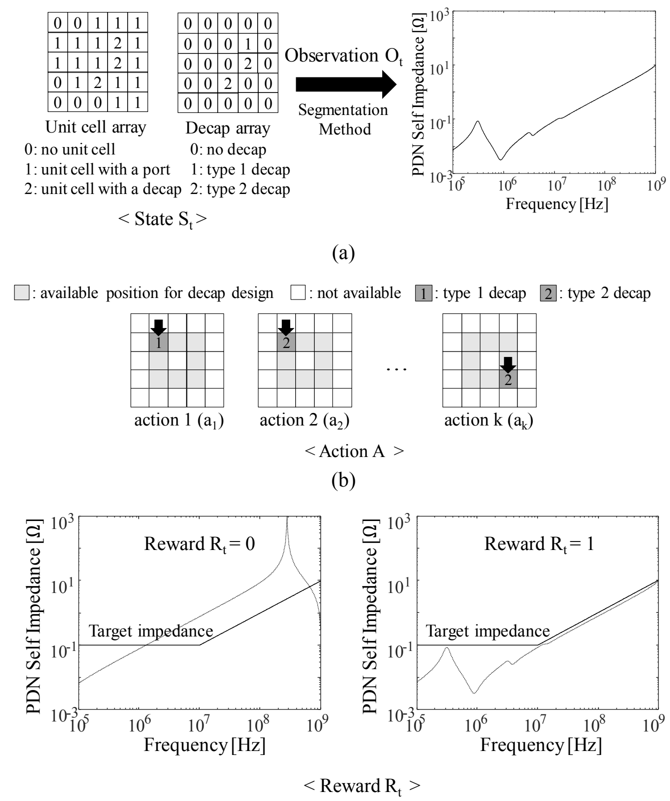

Fig. 3. Definition of the parameters in decap design problem. (a) State S_t and Observation O_t. (b) Action A. (c) Reward R_t.

B. Q-learning algorithm in the proposed method

Algorithm: Q-learning for the proposed method

Set Q(s,a) $\forall s \in S, a \in A$ *arbitrarily*
for *each training episode do*
  $S_{t=0}$ ← *initial state (initial PDN)*
  for *each time step of the episode* ***do***
    Choose A_t at S_t using policy based on Q (ϵ-greedy)
    Take action A_t, get R_t and S_{t+1}
    $Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha[R_t + \gamma \max_a Q(S_{t+1}, a) - Q(S_t, A_t)]$
Until *Z_{11} at IC of S_t satisfies the Z_{target}*

Optimal Q-value can be obtained by applying the Q-learning algorithm in the proposed method. For each time step of each training episode, the decap design agent choose A_t, i.e. assigns a certain type of the decap on a certain position, and obtains the reward R_t based on the comparison between current Z_{11} at IC and the Z_{target}. After that, the agent updates the Q-value as in the following equation:

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha[R_t + \gamma \max_a Q(S_{t+1}, a) - Q(S_t, A_t)]. \quad (1)$$

The discount factor (γ) is set to 0.8, which is a value between 0 and 1, to consider the number of decaps used for decap design. An optimal Q-value q_π^* can be obtained with numerous repetition of training episodes. Finally, the decap design guideline represented by optimal policy $\pi^*(a|s)$ can be obtained as the following equation:

$$\pi^*(a|s) = \begin{cases} 1, & if\ a = argmax_a\, q_\pi^*(a|s) \\ 0, & otherwise \end{cases}. \quad (2)$$

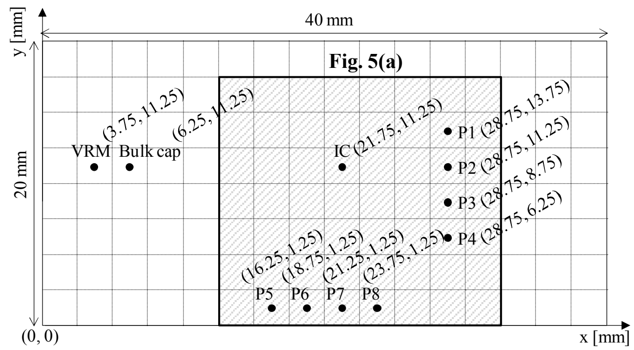

Fig. 4. Configuration of the test on-board PDN with total 11 ports assigned.

TABLE I. PARAMETERS OF AVAILABLE DECAPS

Type	Parameters		
	Capacitance [μF]	***ESL [nH]***	***ESR [mΩ]***
1	10	0.5	5
2	1	0.5	9
3	0.1	0.2	3

III. VERIFICATION OF THE PROPOSED METHOD

The proposed method is applied to the test on-board PDN. In Fig. 4, the configuration of the test on-board PDN is represented. The dimension of the test on-board PDN is 40 mm by 20 mm and consisted of 4 layers where second and third layers are ground and power plane respectively. On the top layer, total of 11 ports are assigned. One port is assigned to each of a voltage regulator module (VRM), a bulk capacitor, and the IC. The remaining 8 ports (P1-P8) are assigned for decap design. The unit cell of the power/ground plane of the PDN, VRM, and bulk capacitor are modeled as RLGC lumped model and connected by using segmentation method [4]. Three types of decaps are assumed to be available and also modeled as lumped model based on S-parameter models. Modeled parameters of the available decaps are listed in Table I.

By applying the proposed method, 37 optimal decap designs are provided where 4 decaps are used each. Fig. 5 shows one of the optimal decap design results and one of the overdesigned decap designs, and corresponding Z_{11} of each PDN compared to the Z_{target}. As shown in Fig. 5(a), for the optimal decap design, type 2 and type 3 decaps are assigned at port P1 and port P3 respectively, and type 1 decaps are assigned at port P5 and P8. As shown in Fig. 5(b), Z_{11} of initial PDN (state #0) unsatisfied the Z_{target} in the frequency range from 1.3 MHz to 657.4 MHz, however, Z_{11} of the decap design optimized PDN (state #2607) satisfied the Z_{target} because assigned decaps suppressed the PDN impedance. In addition, the proposed method provided the optimal decap design with the minimum number of decap used. Z_{11} of decap overdesigned PDN (state #17585) also satisfied the Z_{target}. However, total 5 decaps are assigned as shown in Fig. 5 (a), which is an increase of one more unnecessary decap resulting in the additional cost.

IV. CONCLUSION

In this paper, we firstly proposed the reinforcement learning-based on-board optimal decoupling capacitor design method. For verification, the proposed method is applied to the test on-board PDN and successfully provided 37 optimal decap designs. One of the provided optimal designs is represented in this paper. The corresponding PDN self impedance at the IC satisfied the required Z_{target} with the minimum number of decaps. Moreover, this result is compared to the overdesigned decap design. The proposed method can be extended to not only the on-package, interposer, chip level, but also hierarchical decap design.

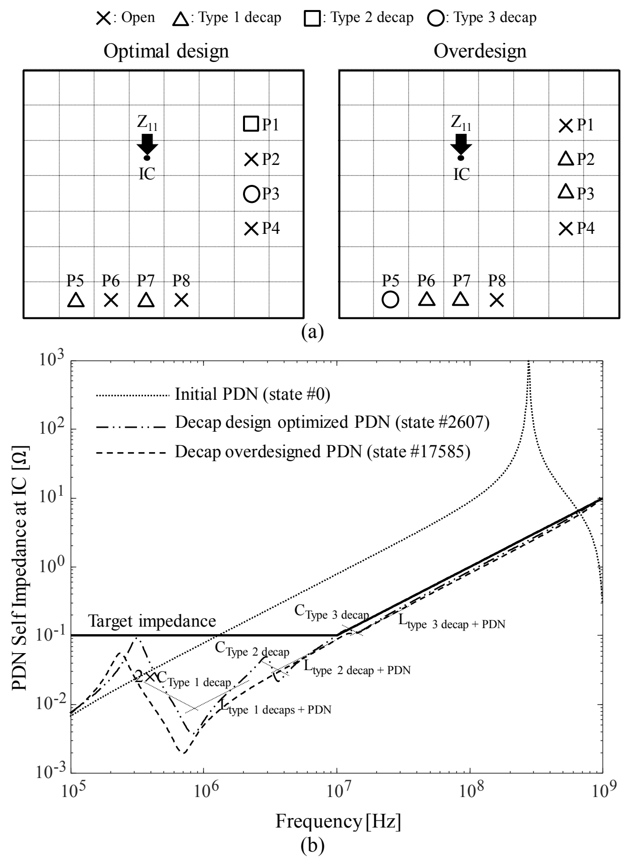

Fig. 5. (a) Optimal and overdesigned decap design result. (b) Self impedance at IC of PDN with optimal and overdesigned decap design.

ACKNOWLEDGMENT

This work was supported by International Collaborative R&D Program (funded by the Ministry of Trade, Industry and Energy (MKE, Korea) [N0000899, Glass interposer based RF FEM for Next Generation Mobile Smart Phone]. We would like to acknowledge the technical support from ANSYS Korea.

REFERENCES

[1] Yang Liu, Yu-Zhang Yuan, et al, "Decoupling capacitors selection algorithm based on maximum anti-resonance points and quality factor of capacitor", Electronics Letters, 8th January 2015, vol. 51 No. 1 pp.90-92.

[2] G. Han, "Simple and Fast Method of On-board Decoupling Capacitor Selection and Placement", 2017 IEEE Electrical Design of Advanced Packaging and Systems Symposium (EDAPS), 2017, pp. 1-3

[3] C. J. Watkins and P. Dayan, "Technical note: Q learning", in *Mach. Learn*, vol. 8, pp. 279-292, 1992.

[4] K.Kim *et al.*, "Modeling and Analysis of a Power Distribution Network in TSV-Based 3-D Memory IC Including P/G TSVs, On-Chip Decoupling Capacitors, and Silicon Substrate Effects," in *IEEE Transactions on Components, Packaging and Manufacturing Technology,* vol. 2 no. 12, pp. 2057-2070, Dec. 2012