We construct the Hayden-Preskill protocol by using a system of spin-1/2 particles and demonstrate information flows of this system which can mimic black holes. We first define an analogous black hole $A$ as a collection of such particles. Second, we take the particles from inside to outside the black hole to define an analogous system of Hawking radiation $B$ as outside particles. When the black hole and the radiation have the maximum entanglement at the Page time, we take an entangled pair system $C$ and $D$. The particles of $C$ fall into the black hole while their counterparts of $D$ remain outside. If we assume rapid mixing of the particle states in the black hole $A \cup C$, can the information of $C$ rapidly escape from the black hole like a mirror? We numerically show that if we turn on the rapid mixing in the black hole, the original information of $C$ rapidly escapes from the black hole to outside in the form of the mutual information between $B$ and $D$. On the other hand, if the mixing between $A$ and $C$ is not enough, the information escapes slowly. Hence, we explicitly demonstrate the original conjecture of Hayden and Preskill. We emphasize that enough mixing is an essential condition to make the Hayden-Preskill protocol functionally work.

Keywords: Black hole information loss problem, Hayden-Preskill protocol, Entanglement entropy

I. INTRODUCTION

Constructing toy models reflecting essential features of quantum mechanics can be useful for tackling the information loss problem of black holes [1,2]. If we present a black hole system as a collection of spin-1/2 particles, we can describe the information flows between the parts of the system straightforwardly [3–7]. Generally, we define information as the difference between the Boltzmann entropy and the entanglement entropy, and it measures the bias from the exact thermal state [8]. As a simple example, we consider a black hole as a bipartite system consisting of a black hole and radiation. In this case, one can fully understand the information flow by calculating the information of the black hole, the information of the radiation, and the mutual information between the black hole and the radiation. The total sum of information is preserved [9]. If we further assume that the states are pure and random, the information of radiation is then negligible by the time the initial Boltzmann entropy of the black hole has decreased to half its value. We call this moment the Page time [3–6].

If the Boltzmann entropy is proportional to the areal radius, then the information should be attached by Hawking radiation, because the black hole will still be semi-classical around the Page time [10]. However, if a Hawking particle carries information, then it causes more
severe problems [11–16]. That some researchers consider various kinds of remnants that carry all information is not surprising [18]. The exact role of the remnant in terms of entanglements and information can be well investigated by using the spin-1/2 toy model [6], where now we need to introduce three parts: the black hole, the radiation, and the remnant. From this toy model, we can obtain several interesting results; for example, the information locking scenario cannot be realized unless the Boltzmann entropy of the final stage of the black hole is very huge [19].

The tension between semi-classical gravity and the unitarity of quantum mechanics can be demonstrated by using a spin-1/2 toy model [7]. The idea that Hawking radiation is generated at the pair creation of two particles around the horizon is reasonable. Note that these two particles are separable from the black hole and the radiation that exist before the pair creation. The number of total states must increase linearly as the pair particles are added. Hence, an annihilation process should take place inside the black hole to prevent this situation. To make this process unitarily, the infalling antiparticle should find a partner particle to be annihilated inside the black hole. This new antiparticle-particle pair should form a new separable state. Interestingly, from the calculation of the toy model, one can show that such a unitary annihilation process is impossible after the Page time [7]. One possible interpretation is that the Hawking particle emission process is not equivalent to the creation of separable pair particles, which violates the no drama condition near the horizon [16,17] which can be potentially very harmful [20–23]. Another interpretation is that the number of states inside a black hole should linearly increase, at least after the Page time [7], although this conclusion is also unsatisfactory [24].

In the sense that the lessons from the spin-1/2 particles show the difficulty of the information loss paradox, they are fruitful and educative. In this context, we will demonstrate one other important idea of the information theoretical issue of the black hole evaporation, which is known by the Hayden-Preskill protocol [25]. According to Hayden and Preskill, a small bit of information, which is inserted to the black hole after the Page time, will be rapidly emitted from the black hole like a mirror reflection. Here, the rapidness is related to the time scale of sufficient randomization or scrambling. The time scale of scrambling is believed to be marginally consistent with the black hole complementarity principle. However, later, one can show that such a marginal bound for black hole complementarity can breakdown [15]. Moreover, if the Hayden-Preskill protocol is true, one can show that the Horowitz-Maldacena proposal [26], which is a candidate resolution of the information loss problem, may still be an unsatisfactory idea [27,28]. On the other hand, if the Hayden-Preskill protocol is not true, some of the interesting previous studies will lose their own grounds.

Therefore, the explicit demonstration of the Hayden-Preskill protocol is an important and interesting issue that we need to investigate. In Section II, we first illustrate the original version of the Hayden-Preskill protocol and second show the model with spin-1/2 particles. In Section III, we show several numerical experiments and confirm that the original Hayden-Preskill conjecture functionally works as long as we introduce a rapid mixing. Finally, in Section IV, we critically review this paper and discuss possible future applications.

II. DEMONSTRATING THE HAYDEN-PRESKILL PROTOCOL

In this section, we first summarize the original argument and theoretical justification of Hayden and Preskill [25]. After this, we illustrate a way to demonstrate the Hayden-Preskill argument by using spin-1/2 particles.

1. Review of the Hayden-Preskill Protocol

Hayden and Preskill [25] considered a situation in which a black hole first exists after the Page time. The total system is the black hole (say, \(A\)) and radiation (say, \(B\)), where the number of states for \(A\) is smaller than that for \(B\). Next, one puts a small bit of information, say \(C\), into the black hole. Finally, one tests when the infalling information will escape from the black hole.

The next natural question is how to confirm that the emitted particle carries information about \(C\). Hayden and Preskill introduce a particle \(D\), which forms a maximally entangled pair with \(C\). While \(C\) falls into the black hole, \(D\) remains outside. Because \(C\) and \(D\) are maximally entangled, one can check whether the emitted information is about \(C\) or not, by comparing the state with \(D\) and the radiation. More rigorously, one can prove that the fidelity about \(D\) is almost perfectly secured if the system is assumed to be random. Hayden and Preskill named this behavior black holes as mirrors.

A key feature of the Hayden-Preskill protocol is the randomizing process applied to the black hole \(A\) and the infalling matter \(C\) after the Page time. However, time is needed to randomize \(A\) and \(C\), and time might have a fundamental limitation. Based on simple arguments, they proposed that the minimum time scale is about \(\sim T^{-1} \log S\), where \(T\) is the temperature and \(S\) is the entropy. In terms of the black hole mass, it becomes \(\sim M \log M\), which they call the scrambling time.

When the paper of Hayden and Preskill was published, they considered this behavior in the context of black hole complementarity [10], which says that at least two complementary observers exist, where one is an asymptotic observer and the other is an infalling observer. According to the asymptotic observer, the black hole is a kind of membrane, and all infalling information attached to the
horizon is thermalized and emitted via Hawking radiation. Because all the processes are causally connected, the asymptotic observer sees no information loss. However, if the information is emitted too rapidly, the asymptotic observer can compare the information with the infalling observer, which can cause a severe inconsistency [29]. Hayden and Preskill reported that the minimum time scale is $\sim M \log M$, which is marginally consistent with the black hole complementarity principle.

However, by increasing the number of matter fields that contribute to Hawking radiation, one can explicitly show that this marginal bound can be irrelevant within the semi-classical regime. This criticism is consistent with [16,17]. It strongly implies that now the original motivation of the Hayden-Preskill protocol is lost. If no membrane exists at the event horizon, then Hawking radiation from the black hole is not equivalent to particles taken directly from $A \cup C$ [7]. Rather, we should introduce a separable separable anti-particle anti-particle pair generated around the horizon. While the outgoing particle becomes Hawking radiation, the anti-particle falls into the black hole. Then can the Hayden-Preskill protocol still work even without using black hole complementarity or the membrane paradigm [30]? This goes beyond the scope of this paper, but we will come back to this question later.

2. Modeling by Spin-1/2 Particles

In order to demonstrate the Hayden-Preskill protocol using spin-1/2 particles, we define the following four subsystems:

- 1. Black hole: $A$, where the number of particles is $N_A$ and the number of states is $a = 2^{N_A}$.
- 2. Radiation: $B$, where the number of particles is $N_B$ and the number of states is $b = 2^{N_B}$.
- 3. Infalling information: $C$, where the number of particles is $N_C$ and the number of states is $c = 2^{N_C}$.
- 4. Reference of $C$: $D$, where the number of particles is $N_D$ and the number of states is $d = 2^{N_D}$.

A. Black hole and radiation: a bipartite system

As an initial condition, we assume that $A$ and $B$ are randomly mixed and that $N_A + N_B = \text{const}$. In the beginning, $N_B = 0$ and as time goes on $N_B$ linearly increases. The quantum state is

$$|\psi_{A \cup B}\rangle = \sum_{i_1, ..., i_{N_A+N_B}} c_{i_1, ..., i_{N_A+N_B}} |i_1, ..., i_{N_A+N_B}\rangle,$$

where $i_j = 1, 2$ (up or down), $|i_1, ..., i_{N_A+N_B}\rangle$ are an orthonormal basis, and $c_{i_1, ..., i_{N_A+N_B}}$ are complex numbers that satisfy the orthonormalization condition. By introducing random numbers $c_{i_1, ..., i_{N_A+N_B}}$, we can prepare a random state. By using this state, we can define the density matrix

$$\rho_{A \cup B} = |\psi_{A \cup B}\rangle \langle \psi_{A \cup B}|.$$  \hspace{1cm} (2)

We can further trace-out the degrees of freedom for each part, e.g.,

$$\rho_A = \text{tr}_B \rho_{A \cup B},$$ \hspace{1cm} (3)

$$\rho_B = \text{tr}_A \rho_{A \cup B}.$$ \hspace{1cm} (4)

By using this, we can define the entanglement entropy

$$S(A|B) = -\text{tr} \rho_A \log \rho_A$$ \hspace{1cm} (5)

and $S(A|B) = S(B|A)$ for a pure state.

From these definitions, we can define information for each part as follows:

- Information for the black hole: $I_A \equiv \log a - S(A|B)$.
- Information for the radiation: $I_B \equiv \log b - S(B|A)$.
- Mutual information between $A$ and $B$: $I_{AB} \equiv S(A|B) + S(B|A) - S(A \cup B) = 2S(A|B)$, where the last equality is derived assuming a pure state.

$I_A$ and $I_B$ show bias from the exact thermal state; hence, if $I_A$ or $I_B$ is greater than zero, then in principle, non-trivial information from the black hole or the Hawking radiation can be measured. Also, $I_A + I_B + I_{AB} = \text{const}$ can be shown, if the total number of particles is preserved.

B. Inserting entangled particles

Now, we need to prepare an entangled state of $C$ and $D$ and put $C$ into the black hole. If $C$ and $D$ are maximally entangled, the mutual information between $C$ and $D$ will be the largest. However, practically there is no physical difference, even when we consider a randomly entangled $C$ and $D$. Following the same path as in the previous subsection, we can define

$$|\psi_{C \cup D}\rangle = \sum_{i_1, ..., i_{N_C+N_D}} c_{i_1, ..., i_{N_C+N_D}} |i_1, ..., i_{N_C+N_D}\rangle.$$ \hspace{1cm} (6)

In order to insert $C$ into $A$, we first consider the total state

$$|\psi\rangle = |\psi_{A \cup B}\rangle \otimes |\psi_{C \cup D}\rangle.$$ \hspace{1cm} (7)

On first glimpse, $A \cup B$ and $C \cup D$ are not entangled. However, we can introduce a randomizing interaction between $A$ and $C$, for example, by using the operation $O^{\alpha\beta}$:

$$O^{\alpha\beta} \equiv S_{A \cup C}^{\beta} \otimes I_{B \cup D},$$ \hspace{1cm} (8)
where $\alpha$ and $\beta$ are randomly chosen indices of $A$ and $C$, respectively, and $S_{A\cup C}^{\alpha\beta}$ is the usual swap gate between $A$ and $C$. Repeating many operations, we can randomize the black hole states $A \cup C$.

Finally, we define $A' \equiv A \cup C$. The density matrix has effectively three parts: $A'$, $B$, and $D$. By tracing out degrees of freedom, for this tripartite system, we can again define the information contents as follows:

- Information for the black hole: $I_{A'} \equiv \log ac - S(A'|B \cup D)$.
- Information for the radiation: $I_B \equiv \log b - S(D|A' \cup B)$.
- Information for the reference: $I_D \equiv \log d - S(D|A' \cup C)$.
- Mutual information between $A'$ and $B$: $I_{A'B} \equiv S(A'|B \cup D) + S(B|A' \cup D) - S(A' \cup B|D)$.
- Mutual information between $B$ and $D$: $I_{BD} \equiv S(B|D \cup A') + S(D|B \cup A') - S(B \cup D|A')$.
- Mutual information between $D$ and $A'$: $I_{DA'} \equiv S(D|A' \cup B) + S(A'|D \cup B) - S(D \cup A'|B)$.
- Tripartite information: $I_{A'B'D} \equiv S(A'|B \cup D) + S(B|D \cup A') + S(D|A' \cup B) - S(A' \cup B|D) - S(B \cup D|A') - S(D \cup A'|B) - S(A' \cup B \cup D)$.

Because the total system is a pure state, the sum of all information is preserved and the tripartite information must vanish [6].

As we take particles from $A \cup C$ to $B$, all information will be reduced to $I_B$, $I_D$, or $I_{BD}$. Because $C$ was entangled with $D$, the information about $C$ can be measured from the mutual information with $D$, i.e., $I_{BD}$. Therefore, by measuring $I_{BD}$, we can check whether the information about $C$ is emitted or not.

III. NUMERICAL RESULTS

In this section, we summarize our numerical results.

1. Null Result: Information Flow of a Bipartite System

First, as a demonstration, we report the null result, where there is no $C$ and $D$. In Fig. 1, the total number of particles is 8. Note that, for all simulations in this paper, due to the random number dependence, we repeated 30 times and the results averaged. In addition, we connected dots smoothly by using B-splines to show a clear physical tendency.

As it has already been observed [3–7], the entanglement entropy follows the relation

$$S(A|B) \simeq \log a$$

for $a < b$; if $b > a$, we can rely on the relation $S(A|B) = S(B|A)$. Before the Page time, the infor-
Demonstration of the Hayden-Preskill Protocol via Mutual Information – Jeong-Myeong Bae et al.

Fig. 3. (Color online) Same condition as in Fig. 2, but without a randomization process.

The mutual information of the black hole $I_A$ decreases and the mutual information between the black hole and radiation $I_{AB}$ increases as time goes on while the information of the radiation $I_B$ is negligible. After the Page time, $I_B$ increases while $I_A$ is negligible and $I_{AB}$ decreases. The sum of all information components is preserved, as we expected.

2. Rapid Information Emission after the Page Time

Now, we insert $C$ into $A$ after the Page time. In Fig. 2, we report the result, where $C$ and $D$ are the states of one particle. Initially, the black hole contains 8 particles. When 4 particles come out from the black hole, it reaches the Page time (pink lines in Fig. 2). After the Page time, we insert $C$ into the black hole while the entanglement entropy is decreasing (say, 6 particles escaped)(green lines in Fig. 2). Then, we turn on the mixing between $A$ and $C$ by using the swap gate. After 100 operations, we continue to take out particles to $B$.

Fig. 4. (Color online) $I_{BD}$ for varying number of mixing operation, say 0, 1, 2, 5, and 100, respectively.

Here, one technical note is that we add $C$ and mix it with $A$, where the original particle $C$ will be removed only in the last stage. If it is the black hole system, then the number of particles of $A$ is much larger than that of $C$; hence, probabilistically, the particles of $C$ will not contribute thoughtfully. However, in our toy model, the probability of choosing $C$ is quite high, which will significantly affect the final result and will spoil our interest in checking whether the mutual information rapidly carries the original information or not.

Keeping this in mind, we report the result. The general tendency is consistent with the previous result without $C$ and $D$. The critical observation is that soon after $C$ falls into the black hole, $I_{BD}$ (blue dashed curve of the upper figure of Fig. 2) increases to its maximum value, which is no longer a trivial result. This is due to the randomization between $A$ and $C$. If we turn off the swap gate, then the result is Fig. 3, which clearly shows that no emission via $I_{BD}$ occurs until the last stage. Therefore, with no randomization process, no mirror-like emission of information can occur, as was expected by Hayden and Preskill.

3. Importance of the Sufficient Mixing

By varying the number of swap gate operations, we can see the tendency of the rapid emission via mutual information. Fig. 4 clearly shows a dependence on the number of mixing operations. If it is zero, then no emission via mutual information occurs. If it is only one and, hence, is insufficient, then the emission carries a little bit of information via mutual information, but not fast enough. If the operation is repeated more, then very quickly, the curve becomes saturated. Thus, we can conclude that rapid mixing is significant for the Hayden-Preskill protocol and that the swap gate operation is very effective.
in ensuring the sufficient mixing of entanglements.

Due to limitations of the computation, we could not extend the number of particles more. However, as we increase the number of particles more and develop a better calculation technique, we may see a clearer relation between mixing and information flow. We leave this topic for a future investigation.

IV. CONCLUSION

In this paper, we constructed the Hayden-Preskill protocol by using spin-1/2 particles and checked numerically whether it functionally works or not. To do this, we introduced four parts corresponding to the black hole A, the radiation B, the infalling information C, and its entangled counterpart D. After the Page time, if C is inserted and if A and C are mixed sufficiently, then the information about C rapidly escapes to B in the form of the mutual information $I_{BD}$. Therefore, we numerically demonstrated that the Hayden-Preskill protocol can be realized using our toy system.

One of the remarkable observations is the importance of rapid mixing between A and C. If no mixing occurs, then emission is not rapid. After A and C are rapidly mixed, one needs to take out a particle from $A \cup C$. In real black holes, rapid mixing is not a strange assumption. However, the question is, where is the place of the rapid mixing? If the place is inside the horizon (hence, near the central singularity), then as many authors are claiming, particles cannot be directly taken from $A \cup C$. However, Hawking-like radiation and information flow are well defined in several examples of analog black holes [31]. The application of the Hayden-Preskill protocol to analog black hole models can be an interesting future research topic.

If we can prove experimentally that no firewall exists near the horizon, then either we need to accept the loss of information or an alternative explanation [32–36]. We hope to find a way for this to be done experimentally, at least, in principle.

ACKNOWLEDGMENTS

DY was supported by the National Research Foundation of Korea (Grant No.: 2018R1D1A1B07049126). This work was supported by the Daegu Gyeongbuk Institute of Science and Technology (DGIST) Undergraduate Group Research Project (UGRP) grant.

REFERENCES

[1] S. W. Hawking, Phys. Rev. D 14, 2460 (1976).
[2] S. W. Hawking, Commun. Math. Phys. 43, 199 (1975). [Erratum-ibid. 46, 206 (1976)].
[3] D. N. Page, Phys. Rev. Lett. 71, 1291 (1993).
[4] D. N. Page, Phys. Rev. Lett. 71, 3743 (1993).
[5] D. N. Page, JCAP 1309, 028 (2013).
[6] J. Hwang et al., Class. Quant. Grav. 34, no. 14, 145004 (2017).
[7] J. Hwang, H. Park, D. Yeom and H. Zoe, Korean Phys. Soc. 73, no. 10, 1420 (2018).
[8] S. Lloyd and H. Pagels, Annals Phys. 188, 186 (1988).
[9] A. Alonso-Serrano and M. Visser, Phys. Lett. B 776, 10 (2018).
[10] L. Susskind, L. Thorlacius and J. Uglum, Phys. Rev. D 48, 3743 (1993).
[11] D. Yeom and H. Zoe, Phys. Rev. D 78, 104008 (2008).
[12] S. E. Hong, D. Hwang, E. D. Stewart and D. Yeom, Class. Quant. Grav. 27, 045014 (2010).
[13] D. Yeom, Int. J. Mod. Phys. Conf. Ser. 1, 311 (2011).
[14] P. Chen, Y. C. Ong and D. Yeom, JHEP 1412, 021 (2014).
[15] D. Yeom and H. Zoe, Int. J. Mod. Phys. A 26, 3287 (2011).
[16] A. Almheiri, D. Marolf, J. Polchinski and J. Sully, JHEP 1302, 062 (2013).
[17] A. Almheiri et al., JHEP 1309, 018 (2013).
[18] P. Chen, Y. C. Ong and D. Yeom, Phys. Rept. 603, 1 (2015).
[19] J. Smolin and J. Oppenheim, Phys. Rev. Lett. 96, 081302 (2006).
[20] D. Hwang, B. -H. Lee and D. Yeom, JCAP 1301, 005 (2013).
[21] W. Kim, B. -H. Lee and D. Yeom, JHEP 1305, 060 (2013).
[22] B. -H. Lee and D. Yeom, Nucl. Phys. Proc. Suppl. 246-247, 178 (2014).
[23] P. Chen et al., Phys. Rev. Lett. 116, 161304 (2016).
[24] B. Giddings, Phys. Rev. D 51, 6860 (1995).
[25] P. Hayden and J. Preskill, JHEP 0709, 120 (2007).
[26] G. T. Horowitz and J. M. Maldacena, JHEP 0402, 008 (2004).
[27] S. E. Hong, D. Hwang, D. Yeom and H. Zoe, JHEP 0812, 080 (2008).
[28] D. Yeom and H. Zoe, J. Korean Phys. Soc. 63, 1109 (2013).
[29] L. Susskind and L. Thorlacius, Phys. Rev. D 49, 966 (1994).
[30] K. S. Thorne, R. H. Price and D. A. Macdonald, Black Holes: The Membrane Paradigm (Yale University Press, 1986).
[31] P. Chen and D. Yeom, Phys. Rev. D 96, no. 2, 025016 (2017).

[32] M. Sasaki and D. Yeom, JHEP 1412, 155 (2014).
[33] B. H. Lee, W. Lee and D. Yeom, Phys. Rev. D 92, no. 2, 024027 (2015).
[34] P. Chen, G. Domènech, M. Sasaki and D. Yeom, JCAP 1604, no. 04, 013 (2016).
[35] P. Chen, G. Domènech, M. Sasaki and D. Yeom, JHEP 1707, 134 (2017).
[36] P. Chen, M. Sasaki and D. Yeom, Eur. Phys. J. C 79, 627 (2019).