Exercise Sheet 7: Channel capacities

Capacities of classical channels

Shannon's noisy channel coding theorem states that the capacity of a (noisy) classical channel T is given by the maximum input-output mutual information:

$$C(T) = \max_{X, p_X} I(X : Y)\,,$$

with the input-output mutual information

$$I(X : Y) = H(X) + H(Y) - H(X, Y) = \sum_{x,y} p(x, y) \log \frac{p(x, y)}{p(x)p(y)}.$$

Here X is an input random variable with distribution p_X, and where Y is the random variable describing the output of the channel T with input X. We will determine the capacity of two example channels.

The first channel we treat is the *binary symmetric channel* or *bitflip channel* T_{BSC}. With probability p, the input bit is flipped ($0 \to 1$ and $1 \to 0$), and with probability $1 - p$ it is transmitted without error. We can also formulate this in terms of the following conditional probabilities of the outputs given the inputs:

$$T_{\mathrm{BSC}}\colon \quad \begin{array}{l} \mathbb{P}[0 \,|\, 0] = \mathbb{P}[1 \,|\, 1] = 1 - p, \\ \mathbb{P}[1 \,|\, 0] = \mathbb{P}[0 \,|\, 1] = p. \end{array}$$

4 P. **Exercise 1.**

3 P. (a) Determine the capacity of the binary symmetric channel $C(T_{\mathrm{BSC}})$ in terms of the binary entropy (which is the entropy of a coin/Bernoulli random variable with probability p)

$$H_2(p) := -p \log p - (1 - p) \log(1 - p). \tag{1}$$

For the rest of the sheet, we assume logarithms to be base 2.

Hint: It may be useful to expand $H(Y|X)$ as $\sum_x p(x)H(Y|X = x)$.

Solution

We have

$$\begin{aligned} I(X : Y) &= H(Y) - H(Y|X) \\ &= H(Y) - \sum_x p_X(x) H(Y|X = x) \\ &= H(Y) - \sum_x p(x) H_2(p) \\ &= H(Y) - H_2(p) \\ &\leq 1 - H_2(p). \end{aligned}$$

The second equality uses the hint, the third equality recognizes that upon fixing an input, Y is just a binary random variable (a coin) with probability p, whose entropy thus is $H_2(p)$. The final inequality uses $H(Y) \leq \log 2 = 1$, because we're dealing with a distribution over $\{0, 1\}$.
We can saturate the above inequality by choosing X as a uniformly distributed random variable, in which case it is easy to check that Y is also uniformly distributed. We conclude that $C(T_{\mathrm{BSC}}) = 1 - H_2(p)$.

1 P. (b) The binary symmetric channel T_{BSC} is a classical communication channel. Write down a quantum channel that generalizes it (in the sense that it acts similarly on computational basis states.)

Solution

The qubit channel

$$\rho \mapsto pX\rho X + (1-p)\rho$$

that we discussed on sheet 5 is as desired.

At the output of the binary symmetric channel, it is not possible to detect if an error occurred or not. If we instead got an output "error", every time an error occurred, our lives might get easier. The channel describing this is the so-called *binary erasure channel* T_{E}. Mathematically, we describe it by adding the "error state" e to the possible states of the system (besides 0 and 1). Formally, the binary erasure channel outputs "error" (i.e. the state e) with probability p and with probability $1-p$ transmits the input without error, leading to the following conditional probabilities:

$$T_{\mathrm{E}}: \quad \begin{array}{l} \mathbb{P}[0 \mid 0] = \mathbb{P}[1 \mid 1] = 1-p, \\ \mathbb{P}[e \mid 0] = \mathbb{P}[e \mid 1] = p. \end{array}$$

9 P. Exercise 2. Let us now determine the capacity of T_{E}.

1 P. (a) We first introduce another random variable Z. Namely, we let $Z = \mathbb{1}_E$ be the indicator random variable for the event $E = \{Y = e\}$ that an erasure error occurs. What is $\mathbb{P}[Z=1]$ in terms of p?

Solution

We have that

$$\begin{aligned} \mathbb{P}[Z=1] &= \mathbb{P}[Z=1 \mid X=0]\mathbb{P}[X=0] + \mathbb{P}[Z=1 \mid X=1]\mathbb{P}[X=1] \\ &= \mathbb{P}[Z=1 \mid X=0](1-\mathbb{P}[X=1]) + \mathbb{P}[Z=1 \mid X=1]\mathbb{P}[X=1] \\ &= p(1-\mathbb{P}[X=1]) + p\mathbb{P}[X=1] \\ &= p. \end{aligned}$$

Note that by definition $\mathbb{P}[Z=1] = \mathbb{P}[Y=e]$, so we could have formulated everything in the calculation above in terms of Y instead of Z.

2 P. (b) Given this new random variable Z (note that its complement is $\neg E = \{Y \neq e\}$), we can write

$$H(Y) = H(Y,Z) = H(Z) + H(Y \mid Z), \tag{2}$$

because Z is a deterministic function of Y. Use the above expansion to show that $H(Y) = H_2(p) + (1-p)H_2(\mathbb{P}[X=1])$, where H_2 is the binary entropy defined in Eq. (1).

Hint: Use that $\mathbb{P}[Y=y \mid Y \neq e] = \mathbb{P}[X=y]$.

Solution

As argued in part (a) of this exercise, we have that $\mathbb{P}[Z = 1] = p$. On the other hand, using the given decomposition for $H(Y)$, we have

$$\begin{aligned} H(Y) &= H(Z) + \mathbb{P}[Z = 0]H(Y \mid Z = 0) + \mathbb{P}[Z = 1]H(Y \mid Z = 1) \\ &= H_2(p) + (1 - p)H(Y \mid Y \neq e) + pH(Y \mid Y = e) \\ &= H_2(p) + (1 - p)H_2(\mathbb{P}[X = 1]) + p \cdot 0\,. \end{aligned}$$

Here, we used:

- $H(Y|Y = e) = 0$ since after conditioning on $Y = e$, Y is deterministic and thus doesn't have any entropy.
- $H(Y|Y \neq e) = H(\mathbb{P}[X = 1])$ since if $Y \neq e$, then the case $Y = 1$ occurs with probability $\mathbb{P}[X = 1]$ (namely if and only if $X = 1$) whereas the case $Y = 0$ occurs with probability $1 - \mathbb{P}[X = 1]$ (namely if and only if $X = 0$).

2 P. (c) Compute the conditional entropy $H(Y \mid X)$.

Solution

If we fix a value $x \in \{0, 1\}$ for X, then Y is either equal to x with probability $1 - p$ or equal to e with probability p. Therefore, $H(Y \mid X = x) = H_2(p)$ for both $x \in \{0, 1\}$. From this, we can conclude that $H(Y \mid X) = H_2(p)$.

2 P. (d) Now proceed similarly to Exercise 1 to determine the channel capacity of the erasure channel.

Solution

Now, we need to maximize the mutual information. To this end, we denote $\mathbb{P}[X = 1] = \tau$ and have

$$\begin{aligned} C(T_{\mathrm{E}}) &= \max_{p_X}\{H(Y) - H(Y|X)\} \\ &= \max_{0\leq\tau\leq1} H_2(p) + (1 - p)H_2(\tau) - H_2(p) \\ &= \max_{0\leq\tau\leq1} (1 - p)H_2(\tau) \\ &= 1 - p\,, \end{aligned}$$

where the maximum of $H_2(\tau)$ is achieved for $\tau = 1/2$. (This is e.g. a consequence of Exercise 1 (a) on Sheet 5.)

1 P. (e) Compare the channel capacities of T_{BSC} and T_{E}. Give an approximate numerical value p_0 such that for $p < p_0$ you would rather communicate via T_{E} than via T_{BSC}.

Solution

Using the results derived above, we have

$$C(T_{\mathrm{E}}) - C(T_{\mathrm{BSC}}) = H_2(p) - p\,.$$

So, we should prefer T_{E} over T_{BSC} whenever $H_2(p) > p$. This is the case for $p < p_0 \approx 0.772908$.

1 P. (f) The erasure channel T_{E} is a classical communication channel. Write down a quantum channel that generalizes it (in the sense that it acts similarly on computational basis states.)

Solution

The qubit-to-qutrit channel

$$\rho \mapsto p|e\rangle\langle e| + (1-p)\rho$$

from $\mathbb{C}^2 \to \mathbb{C}^3$ is as desired, where we picked an orthonormal basis for $\mathbb{C}^3$ as $\{|0\rangle, |1\rangle, |e\rangle\}$ (in essence simply relabling $|2\rangle$ as $|e\rangle$).

Classical capacity of a quantum channel

In the lecture, we saw two alternative characterizations of the classical channel capacity of a quantum channel $\mathcal{E}$, which is given by the Holevo-information $\chi(\mathcal{E})$. We want to establish the equivalence of these expressions.

To this end, recall the definition of the quantum mutual information of a bi-partite quantum system in a state ρ_{AB}

$$I(A:B)_{\rho_{AB}} := S(\rho_A) + S(\rho_B) - S(\rho_{AB}). \tag{3}$$

The Holevo information of channel can be defined using the following scheme: Alice encodes the information of a classical random variable X taking values in $\mathcal{X}$ with probability distribution p_X into a quantum state using a set of states $\{\rho_x\}_{x\in\mathcal{X}}$. To keep track of the classical random variable but formulating everything quantum mechanically, we think of Alice keeping a "notebook" of the information she encoded which we can model as storing that information in another register N using an orthogonal basis $\{|x\rangle\}_{x\in\mathcal{X}}$. From this "notebook" register N, the classical information of X can be completely recovered. Altogether, Alice prepares the bi-partite state

$$\rho_{NA} = \sum_x p_X(x)|x\rangle\langle x|_N \otimes \rho_A^x. \tag{4}$$

Then, the state in system A is sent to Bob using the channel $\mathcal{E}$. Thus, we end up with a final state shared between Alice's notebook and Bob

$$\rho_{NB} = \sum_x p_X(x)|x\rangle\langle x|_N \otimes \mathcal{E}[\rho_A^x]_B. \tag{5}$$

We can now ask for the mutual information between the variable X encoded in N and Bob's output of the channel. Analogously to the classical result, maximizing the mutual information over all possible input variables X and encodings yields the capacity of the quantum channel to transmit classical information, i.e.

$$\chi(\mathcal{E}) = \max_{(X,p_X,\{\rho^x\})} I(N:B)_{\rho_{NB}}. \tag{6}$$

4 P. **Exercise 3.** Show that the Holevo information can be equivalently computed as the largest difference between the entropy of the expected output state and the expected entropy of the output state:

$$\chi(\mathcal{E}) = \max_{(X,p_X,\{\rho^x\})} \left\{ S\left(\mathcal{E}\left[\sum_x p_X(x)\rho^x\right]\right) - \sum_x p_X(x) S\left(\mathcal{E}\left[\rho^x\right]\right)\right\}. \tag{7}$$

Hint: The result of Bonus Exercise 1 of sheet 5 will be useful.

Solution

Let's simply call $\sigma_B^x := \mathcal{E}[\rho_A^x]_B$. The marginal states of ρ_{NB} are

$$\rho_N = \mathrm{Tr}_B[\rho_{NB}] = \sum_x p_X(x)|x\rangle\langle x|, \tag{8}$$

$$\rho_B = \mathrm{Tr}_N[\rho_{NB}] = \sum_x p_X(x)\sigma_B^x. \tag{9}$$

For a classical-quantum state, we have seen on Sheet 5 that

$$S(\rho_{NB}) = H(X) + \sum_x p_X(x)S(\sigma_B^x). \tag{10}$$

Thus, the mutual information is

$$I(N:B)_{\rho_{NB}} = S(\rho_N) + S(\rho_B) - S(\rho_{NB}) \tag{11}$$

$$= H(X) + S\left(\sum_x p_X(x)\sigma_B^x\right) - H(X) - \sum_x p_X(x)S(\sigma_B^x), \tag{12}$$

from which the claim follows.

Let's take another perspective on the Holevo information. The quantum relative entropy between two states ρ and σ is defined as

$$D(\rho\|\sigma) = \begin{cases} \mathrm{Tr}[\rho(\log\rho - \log\sigma)] & \text{if } \mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma) \\ +\infty & \text{else} \end{cases}.$$

Here, the support of a Hermitian operator is the orthogonal complement of its kernel or, equivalently, the span of the eigenstates associated its non-zero eigenvalues. The quantum relative entropy plays a crucial role in quantum information theory and is related to many quantum entropic quantities. Here, we prove one such relationship, namely between the quantum relative entropy and the Holevo information.

4 P. **Bonus Exercise 1.** Prove the following equality:

$$\chi(\mathcal{E}) = \max_{(X,p_X,\{\rho^x\})} \mathbb{E}_X\left[D\left(\mathcal{E}[\rho^X] \,\|\, \mathbb{E}_X[\mathcal{E}[\rho^X]]\right)\right].$$

Hint: You first have to show that the relative entropies involved are not infinite.

Solution

By Exercise 3, it suffices to show that for any $(X, p_X, \{\rho^x\})$ we have

$$S\left(\mathcal{E}\left[\sum_x p_X(x)\rho^x\right]\right) - \sum_x p_X(x) S\left(\mathcal{E}\left[\rho^x\right]\right) = \mathbb{E}_X[D(\mathcal{E}[\rho^X]\|\mathbb{E}_X[\mathcal{E}[\rho^X]])] .$$

First, we show that $\operatorname{supp}(\mathcal{E}[\rho^x]) \subseteq \operatorname{supp}\left(\mathbb{E}_X[\mathcal{E}[\rho^X]]\right)$ holds for all x, so we don't have to worry about the relative entropy being infinite. This can be seen as follows: Suppose $|\psi\rangle$ is in the kernel of $\mathbb{E}_X[\mathcal{E}[\rho^X]]$, i.e., $\mathbb{E}_X[\mathcal{E}[\rho^X]]|\psi\rangle = 0$. Then in particular $0 = \langle\psi|\mathbb{E}_X[\mathcal{E}[\rho^X]]|\psi\rangle = \mathbb{E}_X[\langle\psi|\mathcal{E}[\rho^X]|\psi\rangle]$. As $\mathcal{E}[\rho^x]$ is PSD for all x and thus $\langle\psi|\mathcal{E}[\rho^x]|\psi\rangle \geq 0$ for all x, this average being 0 implies that $0 = \langle\psi|\mathcal{E}[\rho^x]|\psi\rangle = \|\sqrt{\mathcal{E}[\rho^x]}|\psi\rangle\|^2$ for all x. Thus, for all x, $|\psi\rangle$ is in the kernel of $\sqrt{\mathcal{E}[\rho^x]}$, which equals the kernel of $\mathcal{E}[\rho^x]$. So, we have shown that $\ker(\mathcal{E}[\rho^x]) \supseteq \ker\left(\mathbb{E}_X[\mathcal{E}[\rho^X]]\right)$ holds for all x. Taking orthogonal complements, we indeed get that $\operatorname{supp}(\mathcal{E}[\rho^x]) \subseteq \operatorname{supp}\left(\mathbb{E}_X[\mathcal{E}[\rho^X]]\right)$ holds for all x.

Now, we don't have to worry about the quantum relative entropy being infinite and we can do the actual computation. We start from the right hand side. :

$$\begin{aligned}
\mathbb{E}_X[D(\mathcal{E}[\rho^X]\|\mathbb{E}_X[\mathcal{E}[\rho^X]])] &= \mathbb{E}_X\left[\operatorname{Tr}\left[\mathcal{E}[\rho^X]\log\left(\mathcal{E}[\rho^X]\right)\right] - \operatorname{Tr}\left[\mathcal{E}[\rho^X]\log\left(\mathbb{E}_X[\mathcal{E}[\rho^X]]\right)\right]\right] \\
&= \mathbb{E}_X\left[-S\left(\mathcal{E}[\rho^X]\right)\right] - \operatorname{Tr}\left[\mathbb{E}_X[\mathcal{E}[\rho^X]]\log\left(\mathbb{E}_X[\mathcal{E}[\rho^X]]\right)\right] \\
&= -\mathbb{E}_X\left[S\left(\mathcal{E}[\rho^X]\right)\right] + S\left(\mathbb{E}_X[\mathcal{E}[\rho^X]]\right) \\
&= S\left(\mathcal{E}\left[\sum_x p_X(x)\rho^x\right]\right) - \sum_x p_X(x) S\left(\mathcal{E}\left[\rho^x\right]\right) .
\end{aligned}$$

The last step used that, because $\mathcal{E}$ is linear, we have $\mathbb{E}_X[\mathcal{E}[\rho^X]] = \mathcal{E}[\mathbb{E}_X[\rho^X]]$. That's exactly what we wanted to show.

Recap

In this exercise, we will also learn to use the three important representations of channels that we have encountered so far – Choi-Jamiołkowski, Kraus, and Stinespring – in concrete examples rather than just abstractly. Recall the qubit noise channels

$$\begin{aligned}
F_\epsilon(\rho) &\coloneqq \epsilon X\rho X + (1-\epsilon)\rho \\
D_\epsilon(\rho) &\coloneqq \epsilon \operatorname{Tr}[\rho]\frac{\mathbb{I}}{2} + (1-\epsilon)\rho
\end{aligned}$$

where $\epsilon \in [0,1]$, from the last sheet.

6 P. **Exercise 4.** We derive different representations for our two example channels (recall Exercise 4 in Exercise sheet 5).

3 P. (a) Give the Choi-Jamiołkowski state, a Kraus representation, and a Stinespring representation for F_1. Explicitly write down the environment system Hilbert space and the Stinespring dilation unitary.

Hint: For the Stinespring dilation, don't just shut up and calculate. Stop and consider: What is the smallest possible additional Hilbert space which you need to make the evolution unitary?

Solution

From our solution to Exercise 4 (a) in Exercise sheet 5, we already know the Choi state:

$$J(F_1) = \frac{1}{2}\sum_{i,j} |i \oplus 1, i\rangle\langle j \oplus 1, j| = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Next, simply plugging $\epsilon = 1$ into the definition of F_ϵ gives us a Kraus representation $F_1(A) = XAX$ with the single (unitary) Kraus operator X. Finally, as this Kraus operator is unitary, F_1 is a unitary channel, so $F_1(A) = XAX$ is already a Stinespring representation with trivial auxiliary/environment system (i.e., $\mathcal{H}_E = \mathbb{C}$) and with Stinespring unitary X.

3 P. (b) Give the Choi-Jamiołkowski state, a Kraus representation, and a Stinespring representation for D_1. Explicitly write down the environment Hilbert space and the Stinespring dilation isometry (no need to complete it to a unitary).

Solution

Now, we consider D_1. Again, we already determined the Choi state in Exercise 4 (a) in Exercise sheet 5, namely

$$J(D_1) = \frac{\mathbb{I} \otimes \mathbb{I}}{d^2} = \frac{\mathbb{I} \otimes \mathbb{I}}{4}.$$

Next, we show that $\{\frac{1}{2}\mathbb{I}, \frac{1}{2}X, \frac{1}{2}Y, \frac{1}{2}Z\}$ is a set of Kraus operators for D_1. To see this, notice that the Bell states $|\Omega\rangle, (X \otimes \mathbb{I})|\Omega\rangle, (Y \otimes \mathbb{I})|\Omega\rangle, (Z \otimes \mathbb{I})|\Omega\rangle$ form an eigenbasis for $J(D_1)$ (simply because they form an ONB), namely we can write

$$J(D_1) = \frac{1}{4}\Big(|\Omega\rangle\langle\Omega| + (X \otimes \mathbb{I})|\Omega\rangle\langle\Omega|(X \otimes \mathbb{I})^\dagger + (Y \otimes \mathbb{I})|\Omega\rangle\langle\Omega|(Y \otimes \mathbb{I})^\dagger + (Z \otimes \mathbb{I})|\Omega\rangle\langle\Omega|(Z \otimes \mathbb{I})^\dagger\Big).$$

From this, since channels and Choi states are isormorphic, we can directly read off the claimed Kraus operators.
An alternative set of Kraus operators for D_1 is $\{\frac{1}{\sqrt{d}}|i\rangle\langle j|\}_{i,j=0,1}$ is a set of Kraus operators for D_1. To prove this, first notice that $\sum_{i,j}(|i\rangle\langle j|)^\dagger|i\rangle\langle j| = d\mathbb{I}$, so we indeed have a valid set of Kraus operators. To show that it actually describes our channel, we compute

$$\sum_{i,j} \frac{1}{\sqrt{d}}|i\rangle\langle j|\rho(\frac{1}{\sqrt{d}}|i\rangle\langle j|)^\dagger = \frac{1}{d}\left(\sum_j \langle j|\rho|j\rangle\right)\sum_i |i\rangle\langle i| = \mathrm{Tr}[\rho]\frac{\mathbb{I}}{d} = D_1(\rho).$$

We can now construct a Stinespring dilation for D_1 from our Kraus representation (we'll take the one in terms of Pauli matrices). The canonical way of getting a Stinespring isometry U from a set of Kraus operators $\{K_i\}_i$ is via $U(|\cdot\rangle \otimes |0\rangle) = \sum_i K_i|\cdot\rangle \otimes |i\rangle$. That is, our environment Hilbert space is $\mathcal{H}_E = \mathbb{C}^4$ and the action of U is described by

$$\begin{aligned} U(|0\rangle \otimes |0\rangle) &= \frac{1}{2}\left(\mathbb{I}|0\rangle \otimes |0\rangle + X|0\rangle \otimes |1\rangle + Y|0\rangle \otimes |2\rangle + Z|0\rangle \otimes |3\rangle\right) \\ &= \frac{1}{2}\left(|0,0\rangle + |1,1\rangle + i|1,2\rangle + |0,3\rangle\right) \equiv |\psi_0\rangle \\ U(|1\rangle \otimes |0\rangle) &= \frac{1}{2}\left(|1,0\rangle + |0,1\rangle - i|0,2\rangle - |1,3\rangle\right) \equiv |\psi_1\rangle. \end{aligned}$$

The action of U on $|\psi\rangle \otimes |0\rangle$ is then described by

$$U(|\psi\rangle \otimes |0\rangle) = \langle\psi|0\rangle|\psi_0\rangle + \langle\psi|1\rangle|\psi_1\rangle.$$

Therefore, after spectral decomposing a general state ρ, we get

$$\begin{aligned} &U(\rho \otimes |0\rangle\langle 0|)U^\dagger \\ &= \langle 0|\rho|0\rangle|\psi_0\rangle\langle\psi_0| + \langle 0|\rho|1\rangle|\psi_0\rangle\langle\psi_1| + \langle 1|\rho|0\rangle|\psi_1\rangle\langle\psi_0| + \langle 1|\rho|1\rangle|\psi_1\rangle\langle\psi_1|. \end{aligned}$$

To double-check that we did the computation correctly, we can check partial traces:

$$\begin{aligned} \mathrm{Tr}_E[|\psi_0\rangle\langle\psi_0|] &= \frac{1}{4}\left(|0\rangle\langle 0| + |1\rangle\langle 1| + |1\rangle\langle 1| + |0\rangle\langle 0|\right) = \frac{\mathbb{I}}{2} = \ldots = \mathrm{Tr}_E[|\psi_1\rangle\langle\psi_1|], \\ \mathrm{Tr}_E[|\psi_0\rangle\langle\psi_1|] &= \frac{1}{4}\left(|0\rangle\langle 1| + |1\rangle\langle 0| - |1\rangle\langle 0| - |0\rangle\langle 1|\right) = 0 = \ldots = \mathrm{Tr}_E[|\psi_1\rangle\langle\psi_0|], \end{aligned}$$

where the ... hide similar computations. So, we get

$$\mathrm{Tr}_E[U(\rho \otimes |0\rangle\langle 0|)U^\dagger] = \langle 0|\rho|0\rangle\frac{\mathbb{I}}{2} + \langle 1|\rho|1\rangle\frac{\mathbb{I}}{2} = \mathrm{Tr}[\rho]\frac{\mathbb{I}}{2} = D_1(\rho).$$

3 P. **Exercise 5.** Explain how to extend your results achieved in Exercise 4 from $\epsilon = 1$ to arbitrary $\epsilon \in [0, 1]$. *Hint: Don't start from scratch. Rather, determine the respective representations for the part of the channel that you are still missing, then appropriately combine that with your results from Exercise 4.*

Solution

Linearity of the Choi-Jamiołkowski isomorphism can be used to determine the Choi states for general $\epsilon \in [0, 1]$ now that we know the Choi states for $\epsilon \in \{0, 1\}$. Concretely, this looks as follows:

$$J(F_\epsilon) = \frac{\epsilon}{2} \sum_{i,j} |i \oplus 1, i\rangle\langle j \oplus 1, j| + (1 - \epsilon)|\Omega\rangle\langle\Omega|\,,$$
$$J(D_\epsilon) = \frac{\epsilon}{d^2} \mathbb{I} \otimes \mathbb{I} + (1 - \epsilon)|\Omega\rangle\langle\Omega|\,.$$

For the Kraus representation, we first observe that the identity channel $A \mapsto A$ has a Kraus representation with the single Kraus operator $\mathbb{I}$. Now, as our channels of interest are convex combinations of channels with known Kraus operators, we can easily construct their Kraus operators by combining the (suitably weighted) Kraus operators of the constituents:

- F_ϵ has Kraus operators $\{\sqrt{\epsilon}X, \sqrt{1 - \epsilon}\mathbb{I}\}$.
- D_ϵ has Kraus operators $\{\sqrt{1 - \epsilon}\mathbb{I}, \frac{\sqrt{\epsilon}}{2}\mathbb{I}, \frac{\sqrt{\epsilon}}{2}X, \frac{\sqrt{\epsilon}}{2}Y, \frac{\sqrt{\epsilon}}{2}Z\}$. If we summarize the two Kraus operator proportional to the identity into a single one, we get the set of Kraus operators $\{\sqrt{\left(\sqrt{1 - \epsilon}\right)^2 + \left(\frac{\sqrt{\epsilon}}{2}\right)^2}\mathbb{I}, \frac{\sqrt{\epsilon}}{2}X, \frac{\sqrt{\epsilon}}{2}Y, \frac{\sqrt{\epsilon}}{2}Z\} = \{\sqrt{1 - \frac{3\epsilon}{4}}\mathbb{I}, \frac{\sqrt{\epsilon}}{2}X, \frac{\sqrt{\epsilon}}{2}Y, \frac{\sqrt{\epsilon}}{2}Z\}$.

For the Stinespring representation, we can now use the same recipe as before ($U(|\cdot\rangle \otimes |0\rangle) = \sum_i K_i|\cdot\rangle \otimes |i\rangle$) to get the Stinespring isometry for general ϵ from the Kraus operator for general ϵ that we just derived.

Total Points: 26 (+4)