Exercise 1. ***Carnot engine as a source of work***

Here we will look at the derivation of the Landauer's bound for the case when we model the work source as a Carnot engine consisting of two baths. But first, let us recap the case with just one bath.

Suppose that the setting consists of two finite-dimensional systems, the target system S and the machine M. The initial state is an uncorrelated one, with the machine in a thermal state of temperature $\beta \geq 0$ and Hamiltonian H_M. Let U a global unitary we act on both systems; we label the final joint state as ρ'_{SM}, and the respective reduced states as ρ'_S and ρ'_M. We will need two more quantities: the final mutual information between systems S and M

$$I(S:M)_{\rho'_{SM}} = S(\rho'_S) + S(\rho'_M) - S(\rho'_{SM}),$$

and the relative entropy of the final state of the machine with respect to its initial state

$$D(\rho'_M||\rho_M) = \mathrm{tr}(\rho'_M \log \rho'_M) - \mathrm{tr}(\rho'_M \log \rho_M).$$

(a) Show that

$$[S(\rho'_S) - S(\rho_S)] + [S(\rho'_M) - S(\rho_M)] = I(S:M)_{\rho'_{SM}} \geq 0.$$

Solution Let us note that since initially S and M form a separable state, and use the fact that applying a unitary to a state does not change its entropy

$$S(\rho'_{SM}) = S(\rho_{SM}) = S(\rho_S) + S(\rho_M).$$

It follows that

$$[S(\rho'_S) - S(\rho_S)] + [S(\rho'_M) - S(\rho_M)] = S(\rho'_S) + S(\rho'_M) - S(\rho'_{SM}) = I(S:M)_{\rho'_{SM}} \geq 0.$$

The last inequality simply follows from the non-negativity property of mutual information.

(b) Using the result from (a), show the equality form of Landauer's principle

$$\beta\, \mathrm{tr}\left(H_M(\rho'_M - \rho_M)\right) - [S(\rho'_S) - S(\rho_S)] = I(S:M)_{\rho'_{SM}} + D(\rho'_M||\rho_M) \geq 0.$$

How does the usual inequality form of Landauer's principle follow?

Solution From (a), it follows that

$$[S(\rho_S) - S(\rho'_S)] + I(S:M)_{\rho'_{SM}} = [S(\rho'_M) - S(\rho_M)].$$

Using the fact that ρ_M is a thermal state at temperature β, we can infer that

$$S(\rho_M) = \beta\, \mathrm{tr}(H_M \rho_M) + \log[\mathrm{tr}(e^{-\beta H_M})].$$

Substituting $S(\rho_M)$ into the equation, we get

$$[S(\rho'_M) - S(\rho_M)] = \beta\, \mathrm{tr}[H_M(\rho_M - \rho'_M)] + S(\rho'_M) + \mathrm{tr}\left(\rho'_M \log\left(\frac{e^{-\beta H_M}}{\mathrm{tr}(e^{-\beta H_M})}\right)\right).$$

Now note that the last two terms on RHS constitute the relative entropy of the final machine state with respect to its initial state:

$$D(\rho'_M||\rho_M) = -S(\rho'_M) - \operatorname{tr}\left(\rho'_M \log\left(\frac{e^{-\beta H_M}}{\operatorname{tr}(e^{-\beta H_M})}\right)\right).$$

This gives us the desired expression

$$[S(\rho_S) - S(\rho'_S)] + I(S:M)_{\rho'_{SM}} = \beta \operatorname{tr}[H_M(\rho_M - \rho'_M)] - D(\rho'_M||\rho_M)$$
$$\beta \operatorname{tr}\left(H_M(\rho'_M - \rho_M)\right) - [S(\rho'_S) - S(\rho_S)] = I(S:M)_{\rho'_{SM}} + D(\rho'_M||\rho_M).$$

Both relative entropy and mutual information are non-negative quantities, so the inequality follows. In particular, this means that

$$[S(\rho'_S) - S(\rho_S)] \geq \beta \operatorname{tr}\left(H_M(\rho_M - \rho'_M)\right),$$

the change in entropy of the system is bounded from below by the change of energy of the machine (up to the inverse temperature β).

Now let us consider an incoherent setting: suppose that the machine is composed of two parts: H coupled to a hot bath of temperature β_H, and R coupled to a cold bath of temperature β_R, $\beta_R \geq \beta_H \geq 0$. The global unitary U is assumed to be energy-preserving

$$[U, H_S + H_H + H_R] = 0.$$

Let us further introduce the following notation:

- *The entropy change on a subsystem:* $\Delta S_\cdot = S(\rho'_\cdot) - S(\rho_\cdot)$;
- *The energy change on a subsystem:* $\Delta E_\cdot = \operatorname{tr}\left(H_\cdot(\rho'_\cdot - \rho_\cdot)\right)$;
- *The free energy of a state ρ at temperature β:* $F^{(\beta)}(\rho) = \operatorname{tr}(H\rho) - \frac{1}{\beta}S(\rho)$ *(its change is labeled as* $\Delta F^{(\beta)}$*);*
- *The Carnot efficiency of a machine with temperatures β_H and β_R:* $\eta = 1 - \frac{\beta_H}{\beta_R}$.

The initial state ρ_{SRH} is uncorrelated between S, C and H.

(c) *Show that the following equality holds:*

$$-\Delta F_S^{(\beta_R)} - \eta \Delta E_H = \frac{1}{\beta_R}\left(\Delta S_S + \Delta S_R + \Delta S_H + D(\rho'_R||\rho_R) + D(\rho'_H||\rho_H)\right) \geq 0.$$

How does this equality relate to the original Landauer's principle?

Hint: The mutual information is always non-negative due to subadditivity of entropy. Apply this observation to the final mutual information between S, C and H:

$$I(S:R:H)_{\rho'_{SRH}} = S(\rho'_S) + S(\rho'_R) + S(\rho'_H) - S(\rho'_{SRH}) \geq 0.$$

Solution Let us consider

$$I(S:R:H)_{\rho'_{SRH}} = S(\rho'_S) + S(\rho'_R) + S(\rho'_H) - S(\rho'_{SRH}) \geq 0.$$

Furthermore, since the initial state is separate, and entropy is invariant under unitary transformations, we have

$$I(S:R:H)_{\rho'_{SRH}} = \Delta S_S + \Delta S_R + \Delta S_H.$$

From (b), we have

$$\Delta S_R = \beta_R \Delta E_R - D(\rho'_R || \rho_R)$$

and

$$\Delta S_H = \beta_H \Delta E_H - D(\rho'_H || \rho_H).$$

That is,

$$I(S:R:H)_{\rho'_{SRH}} = \Delta S_S + \beta_R \Delta E_R - D(\rho'_R || \rho_R) + \beta_H \Delta E_H - D(\rho'_H || \rho_H).$$

Since the unitary is energy-conserving, $\Delta E_S + E_R + \Delta E_H = 0$. Hence,

$$\Delta S_S - \beta_R \Delta E_S + (\beta_R - \beta_H)\Delta E_H = I(S:R:H)_{\rho'_{SRH}} + D(\rho'_R || \rho_R) + D(\rho'_H || \rho_H).$$

Using the free energy, we can rewrite this as

$$\beta_R[F^{(\beta_R)}(\rho_S) - F^{(\beta_R)}(\rho'_S)] - (\beta_R - \beta_H)\Delta E_H = I(S:R:H)_{\rho'_{SRH}} + D(\rho'_R || \rho_R) + D(\rho'_H || \rho_H).$$

Dividing the expression by β_R, and using the non-negativity of relative entropy and mutual information once again, we obtain the desired result

$$-\Delta F_S^{(\beta_R)} - \eta \Delta E_H = \frac{1}{\beta_R}\left(\Delta S_S + \Delta S_R + \Delta S_H + D(\rho'_R || \rho_R) + D(\rho'_H || \rho_H)\right) \geq 0.$$

In particular, this means

$$-\eta \Delta E_H \geq \Delta F_S^{(\beta_R)},$$

the energy extracted from the hot bath is lower bounded by the increase in free energy, weighted with the inverse Carnot efficiency.

To make a more explicit connection to the original derivation of Landauer's principle, we can consider bounding the heat dissipated into the cold bath, rather than that drawn from the hot bath. Substituting $\Delta E_H = -\Delta E_R - \Delta E_S$, we get

$$\Delta S_S - \beta_H \Delta E_S + (\beta_R - \beta_H)\Delta E_R \geq 0,$$

which is exactly the Landauer's limit for an infinitely hot bath $\beta_H \to 0$.

(d) *Now assume that the target system S is actually a cold bath of temperature β_S, $\beta_S > \beta_R > \beta_H$. We seek to extract energy from S (thus cooling it further) by using the heat flowing from H to R. Show that in this case, the second law involves the Carnot efficiency of an absorption refrigerator.*

Solution From (c) we conclude

$$\Delta S_S - \beta_S \Delta E_S + (\beta_S - \beta_H)\Delta E_S + (\beta_R - \beta_H)\Delta E_R \geq 0$$
$$(\beta_S - \beta_H)\Delta E_S + (\beta_R - \beta_H)\Delta E_R \geq \beta_S \Delta F^{(\beta_S)}.$$

Since S is initially in the thermal state, due to the variational principle its free energy cannot decrease after the transfromation, as Gibbs state has the minimal free energy. Hence,

$$-\frac{\Delta E_S}{\Delta E_R} \leq \frac{\beta_R - \beta_H}{\beta_S - \beta_H} = \eta_{\text{absorb. fridge}}$$