2018 IEEE Conference on Decision and Control (CDC)
Miami Beach, FL, USA, Dec. 17-19, 2018

The Second Law of Controlled Linear Stochastic Thermodynamic Systems over a Noiseless Digital Channel

Songhao Yin and Koji Tsumura

Abstract— **It is known that the conventional second law of thermodynamics is not applicable to thermodynamic systems when feedback control is applied to such systems. A generalized form of the second law should be introduced in this case, which contains an additive term that describes the correlation between the microstates and the measurement outcomes. In this study, we consider a situation where a linear stochastic thermodynamic system, which is in contact with a heat bath, is controlled over a noiseless digital channel to evaluate how channel capacity and control performance are interrelated considering the second law of thermodynamics. We show that in this case, the second law of thermodynamics is inclusive of a term that represents channel capacity. We then show that given a fixed value of free energy difference, we can extract a larger amount of work from the system and obtain higher control performance if more channel capacity is used, in the case where an optimal controller and a proper encoder are used in the control system.**

I. INTRODUCTION

The second law of thermodynamics describes the irreversibility of a thermodynamic process. For a microscopic thermodynamic system that is in contact with a constant temperature heat bath, the second law of thermodynamics can be described by the principle of minimum work:

$$\langle W \rangle \geq \Delta F, \tag{1}$$

where $\langle \cdot \rangle$ denotes the stochastic average, W denotes the work done to the system, and ΔF denotes the Helmholtz free energy difference between two different equilibrium states [2]. The equality holds if and only if the process is reversible. From this equation, ΔF can be considered to be the maximum work required to move from an equilibrium state to another equilibrium state in an isothermal process.

However, in the late 19th century, the validity of the traditional second law (1) was questioned for application to a system incorporating measurements and feedback. A thought experiment concerning this is termed as the Maxwell's demon, the simplest model of which is a single molecule heat engine called the Szilard engine [1]. In recent years, it has been recognized that for the entire system including the demon that performs measurements and feedback, the second law is not broken, and it is still being actively studied. For example, Sagawa & Ueda [6] generalized the second law of thermodynamics with a single measurement and feedback as follows:

$$\langle W \rangle \geq \Delta F - k_B T \mathcal{I} \tag{2}$$

An additive term $-k_B T \mathcal{I}$ appears on the right side of the inequality compared to that one without feedback. Here, $\mathcal{I}$ characterizes the mutual information [15] between the microstates and the measurement outcomes. Fujitani & Suzuki [7] generalized the second law for a linear Langevin system, where measurements are performed at many serial discrete times. They showed that in this case, the additive term reduces to the mutual information between the microstates and the measurement outcomes that correspond to the innovation process of the original system.

From the aforementioned, we can interpret the second law for systems with measurements and feedback by introducing the concept of information into thermodynamics. Moreover, in the field of control theory, the evaluation of the relationship between control performance and channel capacity has been actively studied by introducing the concept of information into control theory, in recent years [9]-[13]. Therefore, the main objective of this paper is to investigate the properties of thermodynamic systems from the viewpoint of control theory and information theory, and combine physics with these two fields. To that end, we will consider a situation where a linear stochastic thermodynamic system that is in contact with a heat bath, is controlled over a noiseless digital channel. We show that in this case, the second law of thermodynamics includes a term that contains channel capacity. We then show that given a fixed value of free energy difference, we can extract a larger amount of work from the system and obtain higher control performance if more channel capacity is used, in the case where an optimal controller and a proper encoder are used in the control system.

The organization of this paper is as follows. In Section II, we give the problem formulation. In Section III, we provide the optimal controller design method and give some essential definitions. In Section IV, we give the main results and proofs. In Section V, we give some discussions on the main results. Finally, in Section VI, we give the conclusions and future works.

II. PROBLEM FORMULATION

We consider a control system which consists of a classical thermodynamic system, an encoder, a noiseless digital channel, a decoder and a controller (Fig. 1.). The microstate x_k, channel input a_k, channel output b_k, decoder output y_k, and

This research is supported by Grant-in-Aid for Scientific Research (B)(16H04382), Japan Society for the Promotion of Science.

S. Yin is with the Department of Information Physics & Computing, the University of Tokyo, Japan songhaoshuai@gmail.com

K. Tsumura is with the Department of Information Physics & Computing, the University of Tokyo, Japan tsumura@i.u-tokyo.ac.jp

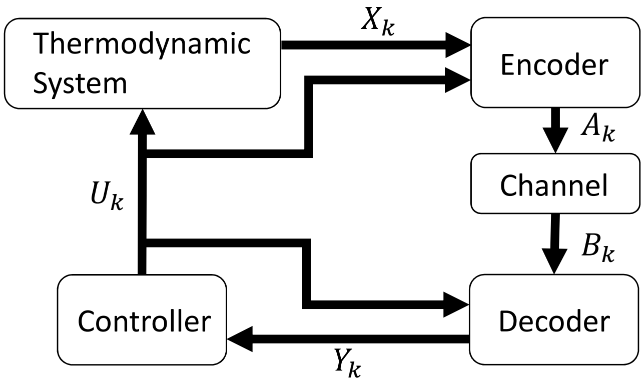

Fig. 1. Control system.

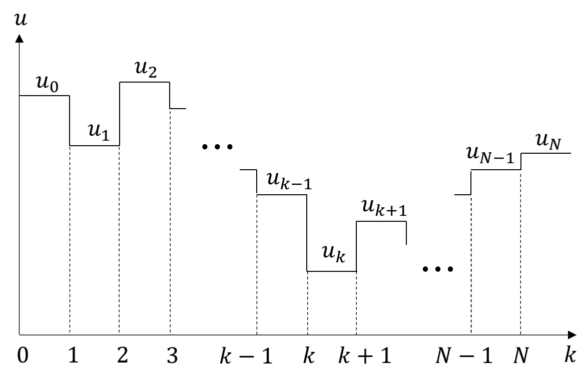

Fig. 2. Control system.

control input u_k are generated in the following order:

$$u_0, x_0, x_1, a_1, b_1, y_1, u_1, \cdots, x_N, a_N, b_N, y_N, u_N \tag{3}$$

Upper case variables like X represent random variables and lower case variables like x represent particular realizations. Let $x_{[1,k]} := (x_1, x_2, \ldots, x_k)$.

A. Thermodynamic System

We assume that the dynamics of the plant, which is a microscopic thermodynamic system that is in contact with a heat bath of temperature T, is given by the following discrete time, stochastic, linear Langevin equation

$$X_{k+1} = FX_k + GU_k + W_k \tag{4}$$

for $k = 0, 1, \cdots, N-1, N \geq 2$, where $X_k \in \mathbb{R}^d$ represents the microstate, $U_k \in \mathbb{R}^m$ represents the control input, and W_k represents the thermal noise. We assume that $\{W_k\}$ is an i.i.d. sequence of random variables with zero mean and covariance K_W. We also assume that (F, G) is controllable.

We fix the input u_0 at time $k = 0$ to be a constant, and keep the plant in a thermal equilibrium state until time $k = 1$ (Fig. 2.). The probability density of x_0 satisfies the canonical distribution under the input u_0. The input value changes from u_k to u_{k+1} at time $k+1$, and maintains a constant value at time interval $[k, k+1)$, while the microscopic state changes from x_k to x_{k+1} in this interval. We denote $H(x_k, u_k)$ as the Hamiltonian of the plant with state x_k under the control input u_k, and denote $H(x_k, u_{k-1})$ as the Hamiltonian just prior to time k. At time $k = N$ we fix the input u_N to a constant. The work $W(x_{[1,N]}, u_{[1,N]})$ done on the plant, the heat $Q(x_{[0,N]}, u_{[0,N-1]})$ absorbed by the plant and the Helmholtz free energy difference $\Delta F(u_0, u_N)$ can be then expressed as follows [4][7]:

$$W(x_{[1,N]}, u_{[1,N]}) = \sum_{k=0}^{N-1} H(x_{k+1}, u_{k+1}) - H(x_{k+1}, u_k), \tag{5}$$

$$Q(x_{[0,N]}, u_{[0,N-1]}) = \sum_{k=0}^{N-1} H(x_{k+1}, u_k) - H(x_k, u_k), \tag{6}$$

$$\Delta F(u_0, u_N) = -k_B T \ln(Z(T, u_N)/Z(T, u_0)), \tag{7}$$

where $Z(T, u_j)$ denotes the partition function. It should be noted that ΔF here is not the difference in the Helmholtz free energy between the initial equilibrium state and the final state at time $k = N$, but the difference between the initial equilibrium state and the equilibrium state with the temperature T under the control protocol u_N, because the final state at time $k = N$ is not in thermal equilibrium.

Incidentally, we can separate the state variable X_k as follows:

$$X_k = \bar{X}_k + \tilde{X}_k \tag{8}$$

$$\bar{X}_{k+1} = F\bar{X}_k + W_k \tag{9}$$

$$\tilde{X}_{k+1} = F\tilde{X}_k + GU_k \tag{10}$$

Eq. (9) is called the innovation process. Assume that $\tilde{x}_0$ is fixed to 0:

$$\tilde{x}_0 = 0 \tag{11}$$

Then, we can use X_k to express the state corresponding to the innovation process $\bar{X}_k$ as follows:

$$\bar{X}_k = X_k - \sum_{i=0}^{k-1} F^{k-i-1} GU_i \tag{12}$$

B. Noiseless Digital Channel

We introduce a noiseless digital channel as the model of communication constraint. The channel input and output alphabets $\mathcal{A}$ and $\mathcal{B}$ are assumed to be the same: $\mathcal{A} = \mathcal{B}$. If the alphabet size is $|\mathcal{A}| = m$, then the rate of the channel or the channel capacity is $R = \ln m$. The channel is noiseless and memoryless, i.e., given a channel input symbol a_k, the channel output b_k satisfies $b_k = a_k$ for all k.

C. Encoder and Decoder

We use the equi-memory expectation predictive encoder and decoder (EMEP encoder and decoder) [12][13], i.e., the encoder is of the form

$$A_k = q\left[\bar{X}_k - E(\bar{X}_k | B_{[1,k-1]})\right], \tag{13}$$

and the decoder is of the form

$$Y_k = \bar{Y}_k + \sum_{i=0}^{k-1} F^{k-i-1} G U_i, \tag{14}$$

where $q[\cdot]$ denotes a quantizer, $E(\cdot)$ represents the expected value of a random variable, and $\bar{Y}_k$ is given by

$$\bar{Y}_k = E(\bar{X}_k | B_{[1,k]}). \tag{15}$$

Note that the microstate is assumed to be fully observed by the encoder. Since both the encoder and decoder have access to the control signals (Fig. 1.), $\bar{X}_k$ can be calculated by the encoder given X_k, and Y_k can be calculated by the decoder given $\bar{Y}_k$. The output of the decoder can be regarded as the estimate of the state of the plant.

Let $\Delta_k = X_k - Y_k$ be the state estimation error. Then, it can be shown that for equi-memory expectation predictive encoders and decoders, $E(\Delta_k) = 0$ and Δ_k is independent of the control actions chosen [12].

D. Cost Function

Consider the following cost function:

$$J_N = E\left(\sum_{k=1}^{N-1} (X_{k+1} - x_d)' \, Q \, (X_{k+1} - x_d) + U_k' S \, U_k\right), \tag{16}$$

where x_d denotes the target state. Let $Q \in \mathbb{R}^{d\times d}$ be positive semi-definite, $S \in \mathbb{R}^{d\times d}$ be positive definite, and $(A, Q^{1/2})$ be observable. The notation X' represents the transposes of X. Given a fixed channel capacity, the value of J_N is determined by the design method of the controller and the quantizer in the encoder. Since the minimization of J_N means to convert the state value X_k to the target value x_d as soon as possible by using control inputs as small as possible, the realization value of J_N can be regarded as an index as the control performance, i.e., a smaller value of J_N means higher control performance. In the following, without loss of generality, we set the value of x_d to 0, i.e., we consider the following cost function for simplicity:

$$J_N = E\left(\sum_{k=1}^{N-1} X_{k+1}' Q X_{k+1} + U_k' S \, U_k\right) \tag{17}$$

E. Research objective

We derive the second law with a term representing the channel capacity, and investigate the relationship between the second law of thermodynamics, channel capacity, and control performance.

III. MINIMIZATION OF THE COST FUNCTION

Under full state observation, it is known that the optimal control law that minimizes (17) is given by the following form [14]:

$$U_k = L_k X_k, \tag{18}$$

$$L_k = -(G' D_{k+1} G + S)^{-1} G' D_{k+1} F, \tag{19}$$

where $D_N = Q, D_k$ $(k = 2, 3, \cdots, N-1)$ is given by

$$D_k = F'(D_{k+1} - D_{k+1} G (G' D_{k+1} G + S)^{-1} G' D_{k+1}) F + Q. \tag{20}$$

In this case, J_N takes the minimum value $\sum_{k=1}^{N-1} \mathrm{tr}(D_{k+1} K_W) + E\left(x_1' D_1 x_1\right)$. This is the well-known LQ stochastic optimal control. Under communication constraints, this result no longer holds. However, we can reduce this problem into a fully-observed LQ problem by introducing a new "fully observed" process [12].

The evaluation function J_N can be rewritten as follows:

$$\begin{aligned} J_N &= E\left(\sum_{k=1}^{N-1} (Y_{k+1} + \Delta_{k+1})' Q (Y_{k+1} + \Delta_{k+1}) + U_k' S \, U_k\right) \\ &= E\left(\sum_{k=1}^{N-1} Y_{k+1}' Q Y_{k+1} + U_k' S \, U_k\right) + E\left(\sum_{k=1}^{N-1} \Delta_{k+1}' Q \Delta_{k+1}\right), \end{aligned} \tag{21}$$

where the second equality holds because $E\left(Y_{k+1}' Q \Delta_{k+1} | b_{[1,k+1]}, u_{[1,k]}\right) = 0$ holds for every $(b_{[1,k+1]}, u_{[1,k]})$. Since Δ_k is independent of the control actions chosen, the second term of (21) is not affected by control inputs. Therefore, the controller should be chosen to minimize the first term of J_N given a fixed channel capacity and encoder.

If we define a new "fully observed" process with the decoder's estimate of the state, Y_k, as the new state, our new system has the dynamics

$$Y_{k+1} = F Y_k + G U_k + \bar{\bar{W}}_k, \tag{22}$$

where $\bar{\bar{W}}_k = E(F\Delta_k + W_k | B_{[1,k+1]})$ can be considered to be the process noise. Since $\{\bar{\bar{W}}_k\}$ are uncorrelated [12], the dynamics (22) with a cost given by the first term of (21) is a fully observed LQ problem and the control inputs can be chosen to be the same as (18) − (20). We refer to the controller given by (18)−(20) as the optimal controller, which is independent of the channel capacity and the design method of the quantizer in the encoder.

We now calculate J_N given the optimal controller. Since Δ_{k+1} can be written as

$$\Delta_{k+1} = F\Delta_k + W_k - \bar{\bar{W}}_k, \tag{23}$$

$K_{\bar{\bar{W}}_k}$, the covariance matrix of $\bar{\bar{W}}_k$, can be written as follows:

$$K_{\bar{\bar{W}}_k} = F \Lambda_k F' + K_W - \Lambda_{k+1} \tag{24}$$

Using (24), J_N can be calculated as follows:

$$\begin{aligned} J_N &= \sum_{k=1}^{N-1} \mathrm{tr}\left(D_{k+1} K_{\bar{\bar{W}}_k}\right) + E\left(\Delta_{k+1}' Q \Delta_{k+1}\right) + E\left(x_1' D_1 x_1\right) \\ &= \sum_{k=1}^{N-1} \mathrm{tr}\left(D_{k+1}\left(F \Lambda_k F' + K_W - \Lambda_{k+1}\right)\right) \\ &\quad + E\left(\Delta_{k+1}' Q \Delta_{k+1}\right) + E\left(x_1' D_1 x_1\right) \\ &= \hat{J}_N + \sum_{k=1}^{N-1} E\left(\Delta_k' M_k \Delta_k\right), \end{aligned} \tag{25}$$

where

$$\hat{J}_N = \sum_{k=1}^{N-1} \mathrm{tr}\left(D_{k+1} K_W\right) + E\left(x_1' D_1 x_1\right), \tag{26}$$

$$M_1 = F' D_2 F, \tag{27}$$

$$M_k = F' D_{k+1} F + Q - D_k \ \ (k = 2, 3, \cdots, N-1) \tag{28}$$

Before moving to the main results, we introduce some definitions that will turn out to be important when discussing the relationship between channel capacity and control performance.

Definition 1: The *estimation error measure* $\Theta_{N,\{M_k\}}$ is defined as

$$\Theta_{N,\{M_k\}} := \sum_{k=1}^{N-1} E\left(\Delta_k' M_k \Delta_k\right). \tag{29}$$

We can easily find that $\Theta_{N,\{M_k\}}$ can be written as the following:

$$\Theta_{N,\{M_k\}} = J_N - \hat{J}_N. \tag{30}$$

Definition 2: We refer to *a proper encoder* as the one that realizes a given value of $\Theta_{N,\{M_k\}}$ (or $J_N - \hat{J}_N$) with the least channel capacity given the optimal controller. We denote the channel capacity in this case as $R_N^{\{M_k\}}(\Theta_{N,\{M_k\}})$ or $R_N^{\{M_k\}}(J_N - \hat{J}_N)$.

Note that $R_N^{\{M_k\}}(\Theta_{N,\{M_k\}})$ is a decreasing function of $\Theta_{N,\{M_k\}}$ and the value goes to infinity as $\Theta_{N,\{M_k\}} \to 0$ (Fig. 3.).

Definition 3: We refer to *the optimal encoder* as the one that minimizes the value of $\Theta_{N,\{M_k\}}$ given the optimal controller and a fixed channel capacity.

Note that the optimal encoder must be a proper encoder, while a proper encoder is not the optimal encoder in general.

IV. MAIN RESULTS

We now present our main results.

Theorem 1: Consider the microscopic thermodynamic system (4) that is in contact with a heat bath of temperature T and is in a thermal equilibrium state at time $k = 0$. If the system is controlled via a noiseless digital channel with capacity R and EMEP encoder and decoder until time $k = N$, then the following inequality holds:

$$\langle W \rangle \geq \Delta F - k_B T (N-1) R \tag{31}$$

Proof 1: We only show the outline of the proof hereafter for the page limitation. We prove this theorem by showing the following formulas in Steps 1 – 3.

$$\boxed{\text{Step 1} \quad \langle W \rangle \geq \Delta F - k_B T I\left(\bar{X}_{[1,N-1]}; \bar{Y}_{[1,N-1]}\right)}$$

Following [7], in the case where the initial states of the forward and backward processes are in equilibrium under u_0 and $u_0^\dagger$, the detailed fluctuation theorem [3]-[5] can be written as

$$e^{\beta(\Delta F - W)} G = \overleftarrow{G}, \tag{32}$$

where

$$G = f_0(x_0) f_{k+1|k}(x_{k+1}|x_k; u_k), \tag{33}$$

$$\overleftarrow{G} = \overleftarrow{f}_0(x_N^*) \overleftarrow{f}_{k|k+1}(x_{k+1}^*|x_k^*; u_k^*). \tag{34}$$

Here, the superscript $*$ denotes the time reversal, $f_0(x_0)$ and $\overleftarrow{f}_0(x_N^*)$ denote the probability density of the initial state of the forward and backward processes respectively, $f_{k+1|k}(x_{k+1}|x_k; u_k)$ denotes the conditional probability density of X_{k+1} given that x_k and u_k are fixed, and $\overleftarrow{f}_{k|k+1}(x_{k+1}^*|x_k^*; u_k^*)$ denotes the conditional probablity density of X_{k+1}^* given that x_k^* and u_k^* are fixed.

From (4), (8), (9) and (11), we can obtain

$$G = f(\bar{x}_{[0,N]}). \tag{35}$$

On the other hand, for the backward process, no estimation is performed and all the actuating signals have already been determined by the values of $\bar{y}_{[1,N-1]}$. Thus, we can rewrite (34) into

$$\overleftarrow{G} = \overleftarrow{f}_0(x_{[0,N]}^*|\bar{y}_{[1,N-1]}). \tag{36}$$

Here, we note that $f(\bar{x}_{[0,N]}|\bar{y}_{[1,N-1]})$ can be rewritten as

$$f\left(\bar{x}_{[0,N]}|\bar{y}_{[1,N-1]}\right) = f\left(\bar{x}_{[0,N]}\right) e^{I\left(\bar{x}_{[1,N-1]}|\bar{y}_{[1,N-1]}\right)}, \tag{37}$$

where

$$I\left(\bar{x}_{[1,N-1]}|\bar{y}_{[1,N-1]}\right) := \ln \frac{f\left(\bar{x}_{[1,N-1]}|\bar{y}_{[1,N-1]}\right)}{f\left(\bar{x}_{[1,N-1]}\right)}. \tag{38}$$

We can thus rewrite (35) as:

$$G = e^{-I\left(\bar{x}_{[1,N-1]}|\bar{y}_{[1,N-1]}\right)} f\left(\bar{x}_{[0,N]}|\bar{y}_{[1,N-1]}\right) \tag{39}$$

Let $p(\cdot)$ denote the probability mass function. From (32), (36) and (39), we obtain

$$\begin{aligned} 1 &= \int dx_{[0,N]} \sum_{\bar{y}_{[1,N-1]}} p\left(\bar{y}_{[1,N-1]}\right) \overleftarrow{G} \\ &= e^{\beta \Delta F} \langle e^{-\beta W - I\left(\bar{x}_{[1,N-1]}|\bar{y}_{[1,N-1]}\right)} \rangle. \end{aligned} \tag{40}$$

Finally, by using Jensen's inequality, we can obtain the inequality in Step 1:

$$\langle W \rangle \geq \Delta F - k_B T I\left(\bar{X}_{[1,N-1]}; \bar{Y}_{[1,N-1]}\right) \tag{41}$$

Here, $I\left(\bar{X}_{[1,N-1]}; \bar{Y}_{[1,N-1]}\right)$ is the mutual information [15][16] between $\bar{X}_{[1,N-1]}$ and $\bar{Y}_{[1,N-1]}$, where $\bar{X}_k$ is a continuous random variable and $\bar{Y}_k$ is a discrete random variable.

$$\boxed{\text{Step 2} \quad I\left(\bar{X}_{[1,N-1]}; \bar{Y}_{[1,N-1]}\right) = H\left(\bar{Y}_{[1,N-1]}\right)}$$

From (13), we can conclude that b_k is uniquely determined given $\bar{x}_k$ and $b_{[1,k-1]}$. We denote this by

$$(\bar{x}_k, b_{[1,k-1]}) \Rightarrow b_k. \tag{42}$$

Now we suppose that $\bar{x}_{[1,k]} \Rightarrow b_{[1,k]}$. From (13) and (42) we obtain

$$\bar{x}_{[1,k+1]} \Rightarrow (b_{[1,k]}, \bar{x}_{k+1}) \Rightarrow (b_{[1,k]}, b_{k+1}) \Rightarrow b_{[1,k+1]}. \tag{43}$$

From the above and the fact that $\bar{x}_1 \Rightarrow b_1$, we can conclude that $\bar{x}_{[1,k]} \Rightarrow b_{[1,k]}$. Furthermore, we obtain from (15) that

$$b_{[1,k]} \Rightarrow \bar{y}_k. \tag{44}$$

From (42), (44) and $\bar{x}_{[1,k]} \Rightarrow b_{[1,k]}$, we have

$$\bar{x}_{[1,k]} \Rightarrow \bar{y}_{[1,k]}. \tag{45}$$

Therefore,

$$H\left(\bar{Y}_{[1,N]}|\bar{X}_{[1,N]}\right) = 0, \tag{46}$$

where $H(\cdot)$ denotes entropy [15]. We then obtain the equation in Step 2 as

$$\begin{aligned} I\left(\bar{X}_{[1,N]}; \bar{Y}_{[1,N]}\right) &= H\left(\bar{Y}_{[1,N]}\right) - H\left(\bar{Y}_{[1,N]}|\bar{X}_{[1,N]}\right) \\ &= H\left(\bar{Y}_{[1,N]}\right). \end{aligned} \tag{47}$$

$\boxed{\text{Step 3} \quad \langle W \rangle \geq \Delta F - k_B T\,(N-1)\,R}$

It is clear from (15) that $b_{[1,k]} \Rightarrow \bar{y}_{[1,k]}$, so we have the following inequality:

$$\begin{aligned} H(B_{[1,N-1]}) &= H(B_{[1,N-1]}, \bar{Y}_{[1,N-1]}) \\ &= H(\bar{Y}_{[1,N-1]}) + H(B_{[1,N-1]}|\bar{Y}_{[1,N-1]}) \\ &\geq H(\bar{Y}_{[1,N-1]}) \end{aligned} \tag{48}$$

Therefore,

$$\begin{aligned} H\left(\bar{Y}_{[1,N-1]}\right) &\leq H\left(B_{[1,N-1]}\right) = H\left(A_{[1,N-1]}\right) \\ &\leq \sum_{k=1}^{N-1} H\left(A_k\right) \leq (N-1)\,R. \end{aligned} \tag{49}$$

From (41), (47) and (49), we thus have

$$\langle W \rangle \geq \Delta F - k_B T\,(N-1)\,R. \tag{50}$$

This ends the proof of the theorem.

□

Corollary 1: Consider the microscopic thermodynamic system (4) that is in contact with a heat bath of temperature T and is in a thermal equilibrium state at time $k = 0$. If the system is controlled via the optimal controller, a proper EMEP encoder & decoder and a noiseless digital channel with capacity $R_N^{\{M_k\}}(J_N - \hat{J}_N)$ until time $k = N$, then the following inequality holds:

$$\langle W \rangle \geq \Delta F - k_B T\,(N-1)\,R_N^{\{M_k\}}(J_N - \hat{J}_N) \tag{51}$$

Proof 2: Obvious from Definition 1, 2 and Theorem 1. □

V. DISCUSSIONS

The inequality in Theorem 1 connects the second law with channel capacity, and shows that as more channel capacity is used, less work needs to be done on the system (in other words, more work can be extracted from the system) given a fixed free energy difference. We stress that Theorem 1 holds whether or not the optimal controller or a proper encoder are used in the control system.

On the other hand, Corollary 1 connects the second law with channel capacity and control performance, which holds in the case where the optimal controller and a proper encoder are used in the control system. Figure 3 is a graph that shows the relationship between the value of $\Theta_{N,\{M_k\}}$ (or $J_N - \hat{J}_N$) and the capacity to realize that value when using a proper encoder, with the horizontal axis be $\Theta_{N,\{M_k\}}$ (or $J_N - \hat{J}_N$) and the vertical axis be $R_N^{\{M_k\}}(\Theta_{N,\{M_k\}})$ (or $R_N^{\{M_k\}}(J_N - \hat{J}_N)$). The value taken by $R_N^{\{M_k\}}(\Theta_{N,\{M_k\}})$ (or $R_N^{\{M_k\}}(J_N - \hat{J}_N)$) is limited to $0, \ln 2, \ln 3, \ldots$ The black squares in the graph corresponds to the cases where the optimal encoders are used. It can be seen from the graph that the smaller the value of $\Theta_{N,\{M_k\}}$ is, i.e., the higher the control performance is, the more capacity is needed to realize that control performance. Whereas, by assuming that the equality in (51) holds, we can obtain the relationship between $R_N^{\{M_k\}}(\Theta_{N,\{M_k\}})$ and $\langle W \rangle$, whose graph is shown in Fig. 4. It can be seen from the graph that given a fixed value of ΔF, a smaller amount of work needs to be done on the system, in other words, we can extract a larger amount of work from the system and obtain higher control performance if more channel capacity is used. In this way, the relationship between the second law, the channel capacity and the control performance is elucidated in the case where the optimal controller and a proper encoder are used in the control system.

Fujitani & Suzuki [7] discussed the case where there are no communication constraints but the measurement output Y_k is mixed with sensor noise V_k, i.e., $Y_k = X_k + V_k$. $\{V_k\}$ is assumed to be an independent and identically distributed (i.i.d.) sequence of random variables with covariance K_V. The second law in this case is as follows [7]:

$$\langle W \rangle \geq \Delta F - k_B T I\left(\bar{X}_{[1,N-1]}; \bar{Y}_{[1,N-1]}\right). \tag{52}$$

Here, $I\left(\bar{X}_{[1,N-1]}; \bar{Y}_{[1,N-1]}\right)$ denotes the mutual information between the microstates and the measurement outputs up to time $N-1$, which corresponds to the innovation process. This result corresponds to inequality (41). In general, $I\left(\bar{X}_{[1,N-1]}; \bar{Y}_{[1,N-1]}\right)$ satisfies the following inequality [15]:

$$\begin{aligned} & I\left(\bar{X}_{[1,N-1]}; \bar{Y}_{[1,N-1]}\right) \\ =\; & h\left(\bar{Y}_{[1,N]}\right) - (N-1)\,h\left(V_k\right) \\ \geq\; & h\left(\bar{Y}_{[1,N]}\right) - \tfrac{1}{2}\,(N-1)\ln\left((2\pi e)^d\,|K_V|\right), \end{aligned} \tag{53}$$

where $h(\cdot)$ denotes the differential entropy and d denotes the dimension of V_k. If $|K_V|$ converges to 0, which means that the measurement is accurate, we can conclude that $I\left(\bar{X}_{[1,N-1]}; \bar{Y}_{[1,N-1]}\right)$ goes to ∞ from the inequality above. This means that we can extract an infinite amount of work from the system if the measurement is perfect.

For the case where the control is performed under communication constraints but there is no sensor noise, if the estimation error measure $\Theta_{N,\{M_k\}}$ converges to 0, $R_N^{\{M_k\}}(\Theta_{N,\{M_k\}})$ goes to ∞. Since $\Theta_{N,\{M_k\}} \to 0$ implies perfect measurements, our result is consistent with that obtained by Fujitani & Suzuki [7].

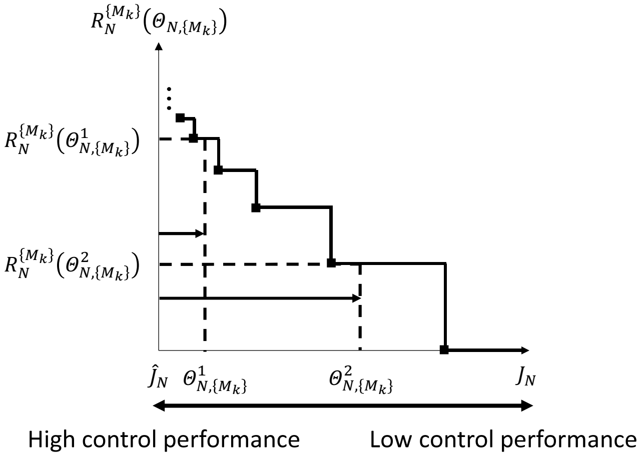

Fig. 3. Graph of $R_N^{\{M_k\}}(\Theta_{N,\{M_k\}})$.

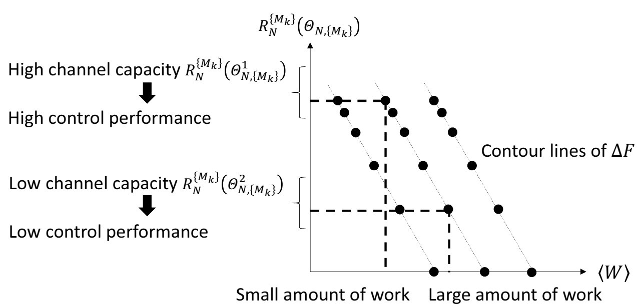

Fig. 4. Relationship between $R_N^{\{M_k\}}(\Theta_{N,\{M_k\}})$ and $\langle W \rangle$.

VI. CONCLUSIONS AND FUTURE WORKS

In this study, we considered a situation in which linear stochastic thermodynamic systems are controlled over a noiseless digital channel, and derived the corresponding second law, which includes a term representing the channel capacity. Moreover, we concluded that given a fixed value of free energy difference, we can extract a larger amount of work from the system and obtain higher control control performance if more channel capacity is used, in the case where an optimal controller and a proper encoder are used in the control system. We will generalize our results for Hamiltonian classical systems and quantum systems.

VII. ACKNOWLEDGMENT

This research is supported by Grant-in-Aid for Scientific Research (B)(16H04382), Japan Society for the Promotion of Science.

REFERENCES

[1] L. Szilard, *Z. Phys. 53*, 1929 [In Germany].

[2] C. Jarzynski, "Nonequilibrium equality for free energy differences", *Phys. Rev. Lett. 78*, 1997.

[3] G. E. Crooks, "Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences", *Phys. Rev. E, 60-3*, 1997.

[4] G. E. Crooks, "Nonequilibrium Measurements of Free Energy Differences for Microscopically Reversible Markovian Systems", *Journal of Statistical Physics*, Vol.90, 1998.

[5] C. Jarzynski, "Hamiltonian derivation of a detailed fluctuatuin theorem", *J. Stat. Phys.*, 2000.

[6] T. Sagawa and M. Ueda, "Generalized Jarzynski Equality under Nonequilibrium Feedback Control", *Phys. Rev.Lett.104*, 090602, 2010.

[7] Y. Fujitani and H. Suzuki, "Jarzynski Equality Modified in the Linear Feedback System", J.Phys.Soc.Jpn.79, 104003,2010.

[8] T. Sagawa and M. Ueda, "Nonequilibrium thermodynamics of feedback control", *Phys. Rev. E 85, 021104*, 2012.

[9] W. Wong and R. Brockett, "Systems with finite communication bandwidth constraints II: Stabilization with limited information feedback", *IEEE Trans. Automat. Contr*, Vol. 44, 1999.

[10] G. Nair and R. Evans, "Mean square stabalisability of stochastic linear systems with data rate constraints", *IEEE Conf. Decision and Control*, 2002.

[11] K. Tsumura and J. Maciejowski, "Stabilizability of SISO control systems under constraints of channel capacities", *IEEE Conf. Decision and Control*, 2003.

[12] S. Tatikonda, A. Sahai and S. Mitter, "Stochastic Linear Control Over a Communication Channel", *IEEE Trans. Automat. Contr*, Vol. 49, No.9, 2004.

[13] S. Tatikonda, "Control under communication constraints", Ph.D. dissertation, Dept. EECS, Mass. Inst. Technol., Cambridge, MA, Aug. 2000.

[14] D. Bertsekas, *Dynamic Programming and Optimal Control*. Belmont, MA: Athena Scientific, 2000.

[15] T. Cover and J. Thomas, "Elements of Information Theory". New York: Wiley, 1991.

[16] B. C. Ross, "Mutual Information between Discrete and Continuous Data Sets", PLOS.ONE, 2014.