Maximum cooling temperature and uniform efficiency criterion for inhomogeneous thermoelectric materials

Zhixi Bian, Hongyun Wang, Qiaoer Zhou, and Ali Shakouri*
Baskin School of Engineering, University of California, Santa Cruz, California 95064, USA
(Received 16 December 2006; published 15 June 2007)

The maximum cooling temperature of a uniform thermoelectric material is limited by its dimensionless figure of merit *ZT*. Inhomogeneous or graded thermoelectric materials are mainly studied when there is a large temperature gradient and the material composition is typically optimized for maximum local *ZT*. We show that this is not the correct optimization for maximizing the cooling temperature. We give the theoretical limit of maximum cooling temperature for an ideal inhomogeneous material. Surprisingly, the optimum Seebeck profile in the device has three sections with distinct characteristics. As a contrast to the local *ZT* optimization, the uniform efficiency criterion is proposed for the design of graded thermoelectric materials in cooling applications. This optimization is applied to the practical Bi_2Te_3 material which is common in thermoelectric applications. Temperature and electrical conductivity dependences of the material properties are taken into account. The graded material is numerically optimized and it achieves a 27% cooling enhancement compared to the best homogeneous material.

DOI: 10.1103/PhysRevB.75.245208 PACS number(s): 72.20.Pa, 05.70.Ce, 85.80.Fi

I. INTRODUCTION

The maximum cooling temperature of a single stage uniform material is limited by its dimensionless figure of merit *ZT*, where *T* is the absolute temperature and the material thermoelectric figure of merit *Z* is a function of Seebeck coefficient *S*, electrical conductivity σ, and thermal conductivity k: $Z=S^2\sigma/k$.[1] The maximum cooling temperature of typical commercial thermoelectric elements ($ZT\sim 1$ at room temperature) is about 70 K. A much larger cooling down from room temperature to cryogenic range is very useful in applications such as low-noise thermal sensors and cameras, superconductor electronic circuits, and microscale low power freezers for biological single cell storage. Since 1990s, nanostructured materials have drawn a lot of attention because their thermal conductivities can be greatly reduced below those of the alloy limits and *ZT* values up to 2.4 have been experimentally observed at room temperature.[2–5] In a parallel direction, a larger cooling temperature beyond that of the *ZT* limit can be pursued through engineering of heat and current flow in three-dimensional configurations.[6] It has been proved that graded thermoelectric materials can achieve much higher cooling temperatures than uniform materials due to the distributed Peltier cooling compensating for the internal Joule heating.[7,8] Essentially, this is about the internal compatibility in a graded thermoelectric cooling element, similar as that in the power generation.[9] However, the global optimization of thermoelectric cooling is even more difficult than that of power generation. This is due to the low efficiency of thermoelectric material. The heat flux through the thermoelectric element stays almost constant in power generation application but increases dramatically from the cold side to the hot side in thermoelectric cooling. In this paper, we first give the theoretical limit of maximum cooling temperature for ideal inhomogeneous materials where the Seebeck coefficient and the electrical conductivity are changing with position but the local *ZT* remains constant. While this is quite a restrictive assumption, the exact analytical solution is quite instructive. It shows that the optimum Seebeck profile has three sections with distinct characteristics. It also provides an upper bound for the maximum cooling. We subsequently propose a criterion, so-called "uniform efficiency," for the design of graded cooling materials. This provides some intuition about the analytic Seebeck profile found earlier. Finally, we use a numerical optimization to maximize cooling in an inhomogeneous state-of-the-art Bi_2Te_3 material. In this case, no simplifying assumptions are made and temperature dependence of material properties and the electrical conductivity dependence of Seebeck coefficient and thermal conductivity are taken into account.

II. MAXIMUM COOLING IN AN IDEAL CASE

Considering a graded thermoelectric material with the thermal conductivity, the Seebeck coefficient, and the electrical conductivity changing with position x, the heat equation at the steady state is

$$\frac{d}{dx}\left[k(x)\frac{dT(x)}{dx}\right]=-\frac{J^2}{\sigma(x)}+JT(x)\frac{dS(x)}{dx}, \qquad (1)$$

where J is the electrical current density and T is the absolute temperature. It can be proved that the cooling performance is independent of the material dimensions. In a practical material, an increase of electrical conductivity is usually accompanied with a decrease of Seebeck coefficient and there is an optimal electrical conductivity which gives the maximum thermoelectric power factor $S^2\sigma$. The thermal conductivity usually increases with the electrical conductivity due to the electronic contribution to heat conduction. For most semiconductor materials, lattice thermal conductivity is dominating and the relative change of thermal conductivity is small when the electrical conductivity changes. In this section, in order to find an analytical solution, we assume that the thermal conductivity $k(x)=C_1$ and the power factor $S(x)^2\sigma(x)=C_2$ are constants in a finite range of electrical conductivity values. This assumption is a little stricter and more practical than that made in Ref. 7. Muller *et al.* change each material parameter independently. Even though the average Seebeck coefficient and electrical conductivity do not change, the local ZT could become larger than the maximum material ZT. Under constant power factor and thermal conductivity assumptions for a graded material with small ZT (<0.2),[8] the maximum cooling temperature with respect to the current density J can be written as

$$\Delta T_{\max} = \frac{1}{2} ZT^2 \frac{\int_0^L S(x)dx \int_0^x S(x')dx'}{\int_0^L dx \int_0^x S^2(x')dx'}. \quad (2)$$

We prove mathematically in the Appendix that the theoretical limit of maximum cooling temperature with respect to the Seebeck profile $S(x)$ is

$$\Delta T_{\max} = \left[1 + \frac{1}{2}\ln\left(\frac{S_L}{S_0}\right)\right]\left(\frac{1}{2}ZT^2\right), \quad (3)$$

where S_0 and S_L represent Seebeck coefficients at the starting and ending positions, respectively. The corresponding optimal Seebeck profile includes three sections,

$$S_{opt}(x) = \begin{cases} S_0, & 0 < x < 1/2 \\ (S_0/2)/(1-x), & 1/2 < x < 1 - S_0/2S_L \\ S_L, & 1 - S_0/2S_L < x < 1, \end{cases} \quad (4)$$

where position x is normalized with the material length L. The optimal Seebeck profile is plotted in Fig. 1(a) for an ideal material, whose thermal conductivity is 125 W/m K, The Seebeck coefficient changes from 150 to 750 μV/K, while its power factor keeps a constant value of 0.0013 W/m K^2, which is a good approximation for a typical silicon material. The theoretical limit [Eq. (3)] represents an increase of 80.5% over $0.5ZT^2$, the maximum cooling temperature of a uniform material. The efficiency of thermoelectric cooling, i.e., the coefficient of performance (COP), is defined as the ratio of the cooling power to the consumed electrical work,

$$\mathrm{COP} = Q_c/W. \quad (5)$$

For a fixed cooling temperature, there is an optimal pairing of heat load and cooling current, which gives the largest COP. For the inhomogeneous material optimized for maximum cooling temperature, the COP could be improved at large cooling temperatures. This is verified in Fig. 1(b), which compares the optimal COP values of uniform and inhomogeneous materials at different cooling temperatures. The intercepts with cooling temperature axis (COP=0) are the maximum cooling temperatures of corresponding materials.

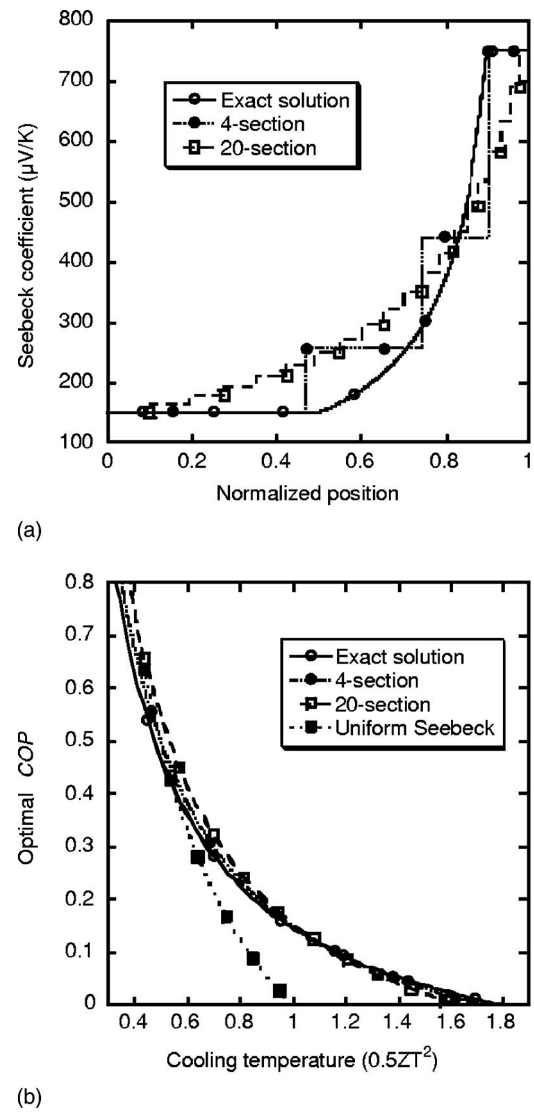

FIG. 1. (a) Three Seebeck profiles of an ideal inhomogeneous material: the exact mathematical solution giving the maximum cooling temperature and the approximate solutions using 4-section and 20-section staircase profiles according to the uniform efficiency criterion. (b) Comparison of the optimal COP of uniform and three inhomogeneous materials at different cooling temperatures.

III. UNIFORM EFFICIENCY CRITERION

Because the optimal graded Seebeck profile defined in Eq. (4) also offers the largest COP when the largest cooling temperature is pursued, it makes sense to find more physical meaning behind the mathematical solutions by thinking about how to improve the cooling efficiency of a graded material. The "global" optimization of the inhomogeneous Seebeck profile is justified in that the heat flux increases dramatically and the "local" efficiency varies from the cold side to the hot side in a uniform material so that different positions cannot reach the optimal operation at the same time. It is reasonable to expect that a uniform distribution of the local efficiency, which resulted from increasing the Seebeck coefficient, could improve the global cooling capacity. In this case, we segment the material into sections and each section is supposed to be homogeneous. The proposed "uniform efficiency criterion" here means that these sections have the same local cooling temperature ΔT and they work individually with the same optimal COP, obeying the relation[1]

$$\mathrm{COP}_{opt} = \frac{T_c}{\Delta T}\frac{(1+ZT_m)^{1/2} - T_h/T_c}{1+(1+ZT_m)^{1/2}}, \quad (6)$$

where T_h, T_c, and T_m are each section's hot end, cold end, and average temperature, respectively. The optimal COP at a given cooling temperature can be achieved only at the optimal cooling current given by

$$\frac{I_{opt}}{I_{\max}} = \frac{\Delta T}{T_c}\frac{1}{(1+ZT_m)^{1/2}-1}, \quad (7)$$

where the maximum cooling current $I_{\max}$ is the ratio of ST_c and the electrical resistance R of this section. The corresponding optimal heat load Q is determined in turn by the unitless equation

$$\frac{\Delta T}{\Delta T_{\max}} = -\left(\frac{I_{opt}}{I_{\max}}\right)^2 + 2\left(\frac{I_{opt}}{I_{\max}}\right) - \frac{Q}{Q_{\max}}, \quad (8)$$

where $Q_{\max}$ is the maximum cooling capacity $S^2T_c^2/2R$. Suppose that the material ZT is small and the total cooling temperature is much smaller than the heat sink temperature. Approximately, T_c, ZT_m, and T_h/T_c can be taken as constants in Eqs. (6) and (7) for different sections. According to Eq. (6), ΔT should be about the same for all the sections if they have the same optimal COP, denoted by ϕ. Since these sections are electrically in series and the electrical currents I_{opt} are the same, the maximum cooling current $I_{\max}$ of all the sections should be the same according to Eq. (7),

$$I_{\max} = \frac{S_i(T_c)_i}{R_i} \approx \frac{S_iTA\sigma_i}{L_i} = C_3, \quad (9)$$

where i is the section index, A the cross-sectional area, L_i the length of section i, and C_3 a constant. Because the electrical work in the preceding sections is added into the heat load of the succeeding section, according to Eq. (5), the heat loads of adjacent sections are related by $Q_{i+1}=Q_i+W_i=Q_i(1+1/\phi)$. Thus, the ratio of the heat loads of adjacent sections is a constant,

$$\frac{Q_{i+1}}{Q_i} = 1 + \frac{1}{\phi} = C_4. \quad (10)$$

Then, it can be seen from Eq. (8) that the ratio of the maximum cooling capacities of adjacent sections is given by the same constant,

$$\frac{(Q_{\max})_{i+1}}{(Q_{\max})_i} = \frac{S_{i+1}^2\sigma_{i+1}L_i}{S_i^2\sigma_iL_{i+1}} = C_4. \quad (11)$$

Using Eqs. (9) and (11) and the constant power factor assumption $S_i^2\sigma_i=C_2$, it can be derived that

$$S_{i+1}/S_i = L_i/L_{i+1} = 1 + 1/\phi = C_4. \quad (12)$$

Suppose that the total length of N cascaded sections is L and the Seebeck coefficients at the cold and the hot sides are, respectively, S_1 and S_N. We determine the local coefficient of performance ϕ, the Seebeck profile $\{S_1, S_2, \ldots, S_N\}$, and the length L_i of each section using Eq. (12) and the following equations derived from Eq. (12):

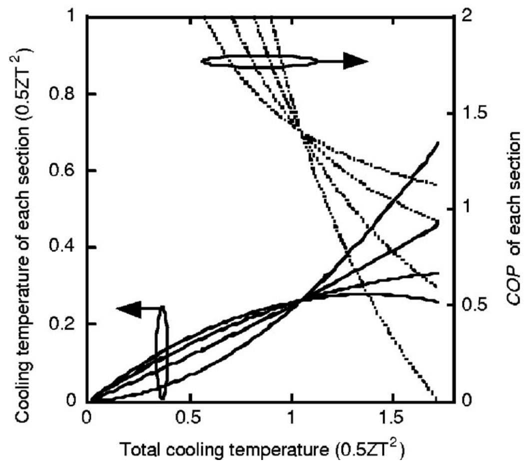

FIG. 2. The cooling temperature and COP of each section for the 4-section material working at its optimal condition for different total cooling temperatures.

$$\frac{S_N}{S_1} = \prod_{i=1}^{N-1}\frac{S_{i+1}}{S_i} = \left(1+\frac{1}{\phi}\right)^{N-1},$$

$$L = \sum_{i=1}^{N} L_i = L_N\sum_{i=1}^{N}\left(1+\frac{1}{\phi}\right)^{i-1} = L_N\phi\left[\left(1+\frac{1}{\phi}\right)^N - 1\right]. \quad (13)$$

The same cooling temperature of each section ΔT can be determined by the uniform local efficiency ϕ from Eq. (6). The total cooling temperature is $\Delta T_{whole}=N\Delta T$. It is easy to find from Eqs. (5) and (10) that the efficiency of the whole material including N sections with the same ΔT and ϕ is

$$\mathrm{COP}_{whole} = \frac{Q_1}{\sum_{i=1}^{N} W_i} = \frac{1}{(1+1/\phi)^N - 1}. \quad (14)$$

The staircase Seebeck profiles of 4-section and 20-section materials and their optimal COP values are plotted in Figs. 1(a) and 1(b), respectively. It can be seen that multiple section optimization according to the criterion of uniform efficiency gives very similar results as the exact solution of global optimization. In fact, when all the local sections are working at their optimal condition with local ΔT and ϕ satisfying Eq. (6), we are not guaranteed that the whole efficiency COP_{whole} is optimal for the total cooling temperature ΔT_{whole} and vice versa. To justify this argument, we scan the cooling current and the heat load to find the optimal working condition of the 4-section material with the maximum COP_{whole} for each ΔT_{whole}. The local cooling temperature and efficiency ϕ of the four sections under the optimal working condition are plotted in Fig. 2. It can be seen that there is only one point of ΔT_{whole}, where both the local cooling temperature ΔT and efficiency ϕ of the four sections are the same and they obey Eq. (6), which means that at this point each section achieves its optimal operation individually as well as the full device. It can also be shown that the optimal total cooling temperature decreases when the section number N increases from $N=4$ (with less than four sections, one cannot define a graded material profile with enough resolution). This makes sense in that each Seebeck profile (section number N in this case) is optimal only for a certain cooling temperature range. These features indicate that the uniform efficiency criterion developed above is suitable for the design of inhomogeneous material optimized for different working conditions and cooling temperatures. Furthermore, choosing a correct number of sections N may push the maximum cooling temperature to its theoretical limit.

IV. PRACTICAL MATERIALS

In a practical material such as Bi_2Te_3 alloy, the thermoelectric power factor always changes with the electrical conductivity. The temperature dependence of material properties should also be taken into account because of its large ZT and cooling temperature. There is no analytical solution of Eq. (1) for the optimal Seebeck profile and the corresponding cooling performance. Usually, the finite element method is used to calculate the total cooling temperature and cooling efficiency of a given material.[10] To facilitate a numerical simulation, we use some approximate formulas in Ref. 11 to describe the relation of material properties and their dependence on the local temperature. We start with a 4-section Seebeck profile designed according to the uniform efficiency criterion using the material properties at 260 K. Then, random variations around the 4-section profile are generated repeatedly to find the optimal case. For each randomly generated profile, the current is scanned to find the maximum cooling temperature and the local material properties are adjusted according to the calculated local temperatures. Several iterations are taken to reach a self-consistent solution. The optimal profiles from different Seebeck ranges are compared to get the best solution. Simulations show that an optimal inhomogeneous Bi_2Te_3 material can achieve a large cooling temperature of 83.9 K, a 27% improvement compared to the optimal uniform material (66.2 K). The cooling efficiency of the optimal inhomogeneous material is much more improved compared to that of the uniform material when the cooling temperature is above 50 K, a typical range in real applications. Figure 3 compares the Seebeck coefficient profile and the local ZT distribution for the optimal uniform and inhomogeneous materials, respectively, when they are operated at their maximum cooling conditions. The slight changes of the Seebeck coefficient and the ZT of the uniform material are due to the temperature dependence of material properties. It is interesting that the optimal profile of the inhomogeneous material still has a shape similar to that of the ideal case [see Fig. 1(a)] after the temperature dependence of material properties and the variation of the thermoelectric power factor are included. It is also prominent that the ZT of the best inhomogeneous material changes a lot in the cooling direction and with a minimum of only 0.47 at the hot end.

Graded thermoelectric materials can be grown using, e.g., epitaxy or sputtering with precisely controlled doping concentration or material composition variation. For highly diffusive material system, a low-temperature growth process is required. High quality Bi_2Te_3 superlattice material has been successfully grown.[12]

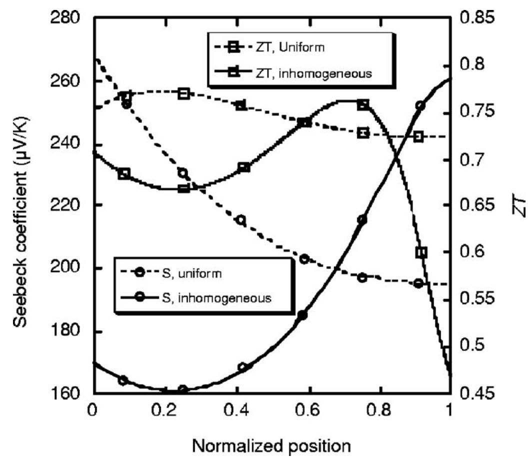

FIG. 3. The Seebeck coefficient profile and the local ZT distribution for the optimal uniform and inhomogeneous Bi_2Te_3 materials.

V. CONCLUSIONS

In summary, we give the maximum cooling of the optimal inhomogeneous material and the corresponding optimal Seebeck coefficient profile for an ideal case. The criterion of uniform efficiency provides some physical insight into the design of inhomogeneous material for different cooling intensities, which can serve as a starting point for a self-consistent numerical optimization of practical materials.

ACKNOWLEDGMENTS

We thank the support by ONR MURI Thermionic Energy Conversion Center. H.W. was partially supported by the NSF.

APPENDIX: NONLINEAR OPTIMIZATION

We want to maximize the functional below with respect to the Seebeck profile $S(x)$,

$$F[S(x)] \equiv \frac{\int_0^1 S(x')\int_0^{x'} S(x)dxdx'}{\int_0^1\int_0^{x'} S^2(x)dxdx'} = \frac{\frac{1}{2}\left[\int_0^1 S(x)dx\right]^2}{\int_0^1 (1-x)S^2(x)dx}. \tag{A1}$$

For a fixed Seebeck profile $S(x)$, the maximum cooling temperature (with respect to the electrical current) is given in Eq. (2) in the text as

$$\Delta T_{\max} = \frac{1}{2}ZT^2F[S(x)]. \tag{A2}$$

In the optimization, the Seebeck profile $S(x)$ is subject to the constraint

$$S_0 \le S(x) \le S_1, \quad x \in [0,1].$$

Below, we will show that the optimal Seebeck profile is given by

$$S_{opt}(x) = \begin{cases} S_0, & 0 \le x \le 1/2 \\ S_0/2(1-x), & 1/2 < x < 1 - S_0/2S_1 \\ S_1, & 1 - S_0/2S_1 \le x \le 1. \end{cases} \tag{A3}$$

It is straightforward to verify that the maximum cooling temperature for $S_{opt}(x)$ is

$$\Delta T_{\max} = \left(\frac{1}{2}ZT^2\right)\left[1 + \frac{1}{2}\ln\left(\frac{S_1}{S_0}\right)\right]. \tag{A4}$$

Now we show step by step the mathematical analysis that leads to $S_{opt}(x)$.

Step 1. The optimal $S(x)$ must be nondecreasing. Suppose that we discretize $S(x)$ as

$$S(x) = y_j \quad \text{for } x \in \left(\frac{j}{N}, \frac{j+1}{N}\right).$$

The numerator and denominator of $F[S(x)]$ are expressed as

$$\frac{1}{2}\left[\int_0^1 S(x)dx\right]^2 = \frac{1}{2}\left(\frac{1}{N}\sum_{j=0}^{N-1} y_j\right)^2,$$

$$\int_0^1 S^2(x)(1-x)dx = \frac{1}{N}\sum_{j=0}^{N-1}\left(1 - \frac{j}{N}\right)y_j^2. \tag{A5}$$

If $S(x)$ is not nondecreasing, then we can find a pair of indices (j,k) such that $j<k$ but $y_j>y_k$. By exchanging the values of y_j and y_k, we have that $(1/2)[(1/N)\Sigma_{j=0}^{N-1}y_j]^2$ is unchanged but $(1/N)\Sigma_{j=0}^{N-1}(1-j/N)y_j^2$ is decreased. It follows that the value of $F[S(x)]$ is increased after the exchange. Therefore, the optimal $S(x)$ must be nondecreasing.

Step 2. The optimal $S(x)$ has three segments,

$$S_{opt}(x) = \begin{cases} S_0, & 0 \le x \le x_0 \\ S_0 < S(x) < S_1, & x_0 < x < x_1 \\ S_1, & x_1 \le x \le 1. \end{cases}$$

This follows directly from that $S_{opt}(x)$ must be nondecreasing (result of step 1).

Step 3. The middle segment of $S_{opt}(x)$ must satisfy

$$S_{opt}(x) = \frac{q}{1-x} \quad \text{for } x \in (x_0,x_1).$$

Let $\Delta S(x)$ be a function that is nonzero only in the middle segment. That is, $\Delta S(x)=0$ in $[0, x_0]$ and in $[x_1, 1]$. Consider a small perturbation to the middle segment of $S_{opt}(x)$,

$$S_{opt}(x) + \varepsilon\Delta S(x).$$

When ε is *small enough*, $S_{opt}(x)+\varepsilon\Delta S(x)$ is between S_0 and S_1 (i.e., satisfying the constraint of the optimization). $S_{opt}(x)$ is the optimal Seebeck profile that implies

$$F[S_{opt}(x) + \varepsilon\Delta S(x)] \le F[S_{opt}(x)], \tag{A6}$$

for ε small enough, which leads to

$$\left.\frac{dF[S_{opt}(x) + \varepsilon\Delta S(x)]}{d\varepsilon}\right|_{\varepsilon=0} = 0. \tag{A7}$$

Restricting our attention to perturbations satisfying $\int_{x_0}^{x_1}\Delta S(x)dx=0$, we have

$$\left.\frac{dF[S_{opt}(x) + \varepsilon\Delta S(x)]}{d\varepsilon}\right|_{\varepsilon=0} = \frac{-2\left[\int_0^1 S_{opt}(x)dx\right]^2}{\left[\int_0^1 (1-x)S_{opt}^2(x)dx\right]^2} \times \int_0^1 (1-x)S_{opt}(x)\Delta S(x)dx,$$

$$\int_{x_0}^{x_1}(1-x)S_{opt}(x)\Delta S(x)dx = 0 \quad \text{for all } \Delta S(x) \text{ satisfying}$$

$$\int_{x_0}^{x_1}\Delta S(x)dx = 0,$$

$$(1-x)S_{opt}(x) = \text{const} \quad \text{for } x \in (x_0,x_1),$$

$$S_{opt}(x) = \frac{q}{1-x} \quad \text{for } x \in (x_0,x_1).$$

Step 4. $S_{opt}(x)$ must be continuous at both x_0 and x_1.

We show the continuity using the method of *proof by contradiction*. Suppose that $S_{opt}(x)$ is discontinuous at x_0. Since $S_{opt}(x)$ is nondecreasing, we have

$$S_0 < \frac{q}{1-x_0} \quad \text{and} \quad \frac{q}{1-x_1} \le S_1,$$

$$(1-x_0)S_0 < q \le (1-x_1)S_1,$$

$$(1-x_0)S_0 - (1-x_1)S_1 < 0. \tag{A8}$$

Let $\Delta S(x)$ be a function that is nonzero only in $[x_0-\delta,x_0]$ and $[x_1,x_1+\delta]$. Specifically,

$$\Delta S(x) = \begin{cases} 1, & x \in [x_0 - \delta, x_0] \\ -1, & x \in [x_1, x_1 + \delta] \\ 0 & \text{otherwise.} \end{cases}$$

Consider a small perturbation to $S_{opt}(x)$,

$$S_{opt}(x) + \varepsilon\Delta S(x).$$

When ε is *positive and small enough*, $S_{opt}(x)+\varepsilon\Delta S(x)$ is between S_0 and S_1 (i.e., satisfying the constraint of the optimization). $S_{opt}(x)$ is the optimal Seebeck profile that implies

$$F[S_{opt}(x) + \varepsilon\Delta S(x)] \le F[S_{opt}(x)],$$

for ε positive and small enough, which leads to

$$\left.\frac{dF[S_{opt}(x)+\varepsilon\Delta S(x)]}{d\varepsilon}\right|_{\varepsilon=0} \leqslant 0.$$

Notice that $\Delta S(x)$ satisfies $\int_{x_0}^{x_1}\Delta S(x)dx=0$. We have

$$\left.\frac{dF[S_{opt}(x)+\varepsilon\Delta S(x)]}{d\varepsilon}\right|_{\varepsilon=0} = \frac{-2\left[\int_0^1 S_{opt}(x)dx\right]^2}{\left[\int_0^1 (1-x)S_{opt}^2(x)dx\right]^2} \times\int_0^1 (1-x)S_{opt}(x)\Delta S(x)dx,$$

$$\int_0^1 (1-x)S_{opt}(x)\Delta S(x)dx \geqslant 0,$$

$$\int_{x_0-\delta}^{x_0} (1-x)S_0dx - \int_{x_1}^{x_1+\delta} (1-x)S_1dx \geqslant 0,$$

$$\left(1-x_0+\frac{\delta}{2}\right)\delta S_0 - \left(1-x_1-\frac{\delta}{2}\right)\delta S_1 \geqslant 0.$$

Dividing by δ and taking the limit as $\delta\to 0$, we obtain

$$(1-x_0)S_0-(1-x_1)S_1 \geqslant 0,$$

which contradicts with Eq. (A8). Therefore, $S_{opt}(x)$ must be continuous at both x_0 and x_1. In other words, $S_{opt}(x)$ has the form

$$S_{opt}(x) = \begin{cases} S_0, & 0\leqslant x\leqslant x_0 \\ q/(1-x), & x_0<x<x_1, \quad x_0=1-q/S_0, \quad x_1=1-q/S_1 \\ S_1, & x_1\leqslant x\leqslant 1. \end{cases} \tag{A9}$$

Step 5. The optimal value of q is $q=S_0/2$,

$$\int_0^1 S_{opt}(x)dx = \int_0^{x_0} S_0dx + \int_{x_0}^{x_1}\frac{q}{1-x}dx + \int_{x_1}^1 S_1dx$$
$$= x_0S_0 + q\log\left(\frac{1-x_0}{1-x_1}\right) + (1-x_1)S_1$$
$$= S_0 - q + q\log\frac{S_1}{S_0} + q = S_0 + q\log\frac{S_1}{S_0}.$$

$$\int_0^1 (1-x)S_{opt}^2(x)dx = \int_0^{x_0}(1-x)S_0^2dx + \int_{x_0}^{x_1}\frac{q^2}{1-x}dx + \int_{x_1}^1 (1-x)S_1^2dx$$
$$= \left(1-\frac{x_0}{2}\right)x_0S_0^2 + q^2\log\left(\frac{1-x_0}{1-x_1}\right) + \frac{(1-x_1)^2}{2}S_1^2$$
$$= \frac{S_0^2-q^2}{2} + q^2\log\frac{S_1}{S_0} + \frac{q^2}{2}$$
$$= \frac{S_0^2}{2} + q^2\log\frac{S_1}{S_0}.$$

Consider the function

$$f(q) \equiv F[S_{opt}(x)] \equiv \frac{\frac{1}{2}\left[\int_0^1 S_{opt}(x)dx\right]^2}{\int_0^1 (1-x)S_{opt}^2(x)dx} = \frac{\left(S_0+q\log\frac{S_1}{S_0}\right)^2}{S_0^2+2q^2\log\frac{S_1}{S_0}}.$$

Taking the derivative with respect to q, we have

$$\frac{df(q)}{dq} = \frac{2\left(S_0+q\log\frac{S_1}{S_0}\right)\log\frac{S_1}{S_0}}{\left(S_0^2+2q^2\log\frac{S_1}{S_0}\right)^2}(S_0^2-2S_0q).$$

The maximum of $f(q)$ is attained at $q=S_0/2$.

In summary, the five steps above have completely determined the optimal solution

$$S_{opt}(x) = \begin{cases} S_0, & 0\leqslant x\leqslant 1/2 \\ S_0/2(1-x), & 1/2<x<1-S_0/2S_1 \\ S_1, & 1-S_0/2S_1\leqslant x\leqslant 1. \end{cases}$$

*Author to whom correspondence should be addressed. Email address: ali@soe.ucsc.edu

[1] D. M. Rowe, *CRC Handbook of Thermoelectrics* (CRC, Boca Raton, FL, 1995), Chap. 3.

[2] M. V. Simkin and G. D. Mahan, Phys. Rev. Lett. **84**, 927 (2000).

[3] G. Chen, A. Narayanaswamy, and C. Dames, Superlattices Microstruct. **35**, 61 (2004).

[4] T. C. Harman, P. J. Taylor, M. P. Walsh, and B. E. LaForge, Science **297**, 2229 (2002).

[5] R. Venkatasubramanian, E. Siivola, T. Colpitts, and B. O'Quinn, Nature (London) **413**, 597 (2001).

[6] Y. Zhang, Z. Bian, and A. Shakouri, in *Proceedings of the 24th International Conference on Thermoelectronics, Clemson, SC* (IEEE, Piscataway, NJ, 2005), p. 248.

[7] E. Muller, S. Walczak, and W. Seifert, Phys. Status Solidi A **203**, 2128 (2006).
[8] Z. Bian and A. Shakouri, Appl. Phys. Lett. **89**, 212101 (2006).
[9] G. J. Snyder and T. S. Ursell, Phys. Rev. Lett. **91**, 148301 (2003).
[10] B. Sherman, R. R. Heikes, and R. W. Ure, J. Appl. Phys. **31**, 1 (1960).
[11] G. D. Mahan, J. Appl. Phys. **70**, 4551 (1991).
[12] R. Venkatasubramanian, T. Colpitts, B. O'Quinn, S. Liu, N. El-Masry, and M. Lamvik, Appl. Phys. Lett. **75**, 1104 (1999).