Macroscopic quantum entanglement of Kondo cloud at finite temperature

S.-S. B. Lee,1 Jinhong Park,1 and H.-S. Sim1

1Department of Physics, Korea Advanced Institute of Science and Technology, Daejeon 305-701, Korea

We propose a variational approach for computing macroscopic entanglement in a many-body mixed state, based on entanglement witness operators, and compute the entanglement of formation (EoF), a mixed-state generalization of entanglement entropy, in single- and two-channel Kondo systems at finite temperature. The thermal suppression of the EoF obeys power-law scaling at low temperature. The scaling exponent is halved from the single- to the two-channel system, which is attributed, using a bosonization method, to the non-Fermi liquid behavior by a Majorana fermion, a “half” of a complex fermion, emerging in the two-channel system. Moreover, the EoF characterizes the size and power-law tail of Kondo screening cloud of the single-channel system.

PACS numbers: 75.20.Hr, 03.67.Mn, 72.15.Qm, 71.10.Hf

Systems of many interacting particles often exhibit unusual macroscopic phenomena at zero temperature. A state-of-the-art concept of understanding their quantum nature is macroscopic entanglement [1, 2], quantum correlation of many particles that cannot be imitated by classical correlations [3]. A popular measure for this purpose is entanglement entropy, which captures entanglement between two macroscopic subsystems. Computation of this measure has revealed new aspects of many-body ground states, including area law [4], topological order [5, 6], and quantum criticality [7, 8].

Generalizing this zero-temperature study is desirable, to explore how the macroscopic entanglement thermally decays or spatially extends. This requires to study a mixed state, in which quantum and classical correlations coexist. At finite temperature, a system is in a probabilistic mixture of energy eigenstates. Its entanglement will reveal quantumness in quantum-to-classical crossover, collective excitations, decoherence, etc. Moreover, entanglement entropy measures the entanglement only between two complementary subsystems in a pure state [1], hence, provides limited information about the spatial extension of macroscopic entanglement. To get more direct information, it is useful to consider, e.g., two distant non-complementary subsystems with changing the distance. The two non-complementary subsystems are described again by a mixed state, after the remainder is traced out of a ground or thermal mixed state.

The computation of macroscopic entanglement in many-body mixed states, however, requires huge costs. For mixed states, entanglement entropy overestimates entanglement, since it cannot distinguish between coexisting quantum and classical correlations. Thus it is generalized [9] into the entanglement of formation (EoF) $\mathcal{E}_F$. EoF quantifies the entanglement between two complementary subsystems A and B of a mixed state $\rho$ as

$$
\mathcal{E}_F(\rho) = \inf_{\rho = \sum_i p_i |\psi_i\rangle \langle \psi_i|} \left( \sum_i p_i \mathcal{E}_E(|\psi_i\rangle) \right).
$$

It is obtained by exploring the possible decompositions of $\rho$ into normalized pure states $|\psi_i\rangle$ with weight $p_i$, and finding the optimal decomposition for which $\sum_i p_i \mathcal{E}_E(|\psi_i\rangle)$ is the lowest. Here, $\mathcal{E}_E(|\psi_i\rangle) \equiv -\text{Tr}(\rho_A \log_2 \rho_A)$ is the entanglement entropy of $|\psi_i\rangle$ between A and B, and $\rho_A = \text{Tr}_B(|\psi_i\rangle \langle \psi_i|)$; the state $\rho_A$ of A is obtained from $|\psi_i\rangle$ by tracing out B. For pure states $|\psi\rangle$, EoF reduces to entanglement entropy, $\mathcal{E}_E(|\psi\rangle) = \mathcal{E}_E(|\psi\rangle)$.

The computation cost of exploring the decompositions is huge even for a small system of a three-qubit full-rank state, equivalent to minimizing a function of 63 ~ 959 variables [10, 12], and it exponentially increases with system size [13, 14]. Most entanglement measures require similar heavy costs [13, 15]; an exception is negativity [16, 20], which however cannot detect some type of entanglement [13, 13].

On the other hand, in the representative many-body problems of Kondo effects [21], the ground states have the entanglement between the Kondo impurity spin and the surrounding conduction electrons, the latter forming Kondo cloud [22, 24]. Naturally, macroscopic entanglement would be a direct tool for characterizing the macroscopic properties of Kondo cloud, the essence of Kondo effects, that cannot be captured by the correlations of a few particles [25, 26]. However, the understanding of the macroscopic entanglement is unsatisfactory due to the computation difficulty mentioned above. For example, it has not been addressed how the entanglement, namely, the macroscopic quantum nature of Kondo cloud, thermally decays. Moreover, the spatial distribution of the cloud, such as the cloud tail, has not been characterized by the entanglement, despite efforts [18, 19, 27, 28].

In this Letter, we propose a variational approach for computing macroscopic entanglement in mixed states, based on entanglement witness operators (EWs) [11, 13, 25, 29, 31], and develop it for single- (1CK) and two-channel Kondo (2CK) systems, using numerical renormalization group (NRG) methods [32, 33]. We compute the EoF $\mathcal{E}_F$ between the Kondo impurity and the electrons located within distance $L$ from the impurity at temperature $T$; see Fig. 1. In addition to the expected crossover around the Kondo temperature $T_{1CK(2CK)}$ of 1CK (2CK), the macroscopic entanglement measured by...
The largest expectation value provides $\mathcal{E}_F(\rho)$.

$$\mathcal{E}_F(\rho) = \sup_{X \in \mathcal{M}_\rho} \text{Tr} X \rho$$

It is equivalent to Eq. (1), and the cost of exploring all operators in $\mathcal{M}_\rho$ is huge. Because of the difficulty in Eqs. (1) and (2), macroscopic entanglement in a thermal many-body state has not been quantified.

In our approach, instead of fully exploring $\mathcal{M}_\rho$, we construct an appropriate variational form of EW, which covers only a small subset of $\mathcal{M}_\rho$ but includes or is close to the optimal EW. Within the form, we find the operator $X_\rho^\text{opt}$ whose expectation value $\text{Tr} X_\rho^\text{opt} \rho$ is the largest.

A lower bound of $\mathcal{E}_F(\rho)$ is obtained as $\text{Tr} X_\rho^\text{opt} \rho$ because $X_\rho^\text{opt}$ is an EW. We obtain an upper bound, based on the duality $\mathbb{I} \leq \mathbb{P}_x \equiv \{|\psi\rangle | \langle \psi| \} = \mathcal{E}_F(|\psi\rangle \langle \psi|)$, if and only if $\text{Tr} X \rho = \mathcal{E}_F(\rho)$ in Eq. (2). We find $\mathbb{P} X_{\rho_{\text{opt}}}^\text{opt}$ and a decomposition of $\rho = \sum_i p_i |\psi_i\rangle \langle \psi_i|$, each $|\psi_i\rangle$ being sufficiently similar to an element of $\mathbb{P} X_{\rho_{\text{opt}}}^\text{opt}$. Then, $\sum_i p_i \mathcal{E}_F(|\psi_i\rangle)$ is an upper bound. The upper and lower bounds are close together (hence to the exact value) when $X_{\rho_{\text{opt}}}^\text{opt}$ is “good”. A good variational form can be constructed for a system at low temperature, considering the entanglement of its ground states and low-energy excitations, as shown below.

**EW in Kondo systems.**—We further develop this approach for Kondo systems, to compute the EoF $\mathcal{E}_F$ between the Kondo impurity and the electrons within distance $L$ from the impurity at temperature $T$; see Fig. 1. We obtain $\rho$, building thermal states by NRG [32, 33] and tracing out the subsystem outside $L$. We develop a way for the latter within NRG [35]. The resulting $\rho$ has rank $\sim 10^4$ too high to exactly obtain $\mathcal{E}_F(\rho)$.

To find an efficient variational form of $X$, we observe the following: The ground state of 1CK and the two-fold degenerate ground states of 2CK are two-qubit Bell states of maximal entanglement ($\mathcal{E}_F = 1$) between impurity spin states ($|\eta = \uparrow, \downarrow\rangle$) and bath states. The energy eigenstates $|E_i\rangle$ with energy $E_i \ll k_BT_{\text{1CK},2\text{CK}}$ have a similar form to two-qubit Bell states $\mathcal{E}_F \lesssim 1$, $|E_i\rangle = b_{\uparrow\downarrow}|\uparrow\rangle|\uparrow\rangle + b_{\downarrow\uparrow}|\downarrow\rangle|\downarrow\rangle$, $b_{\uparrow\downarrow} \approx 1/\sqrt{2}$, $E_{\eta\eta} = \delta_{\eta\eta}$, and $\langle e_{i\eta}|e_{i\eta'}\rangle \approx 0$.

Based on the observation, we propose how to construct $X$ variationally. (i) We decompose the whole Hilbert space into the two-qubit subspaces $\mathcal{H}_i = \{|\uparrow\rangle, |\downarrow\rangle\} \otimes \{|\phi_{\eta\uparrow}\rangle, |\phi_{\eta\downarrow}\rangle\}$, where $|\phi_{\eta\eta}\rangle$ is an orthonormal basis of bath states. $|\phi_{\eta\eta}\rangle$’s are parametrized to be optimized, considering the structure of $\rho$, to make the lower and upper bounds of $\mathcal{E}_F(\rho)$ closer [35]. At $T \ll T_{\text{1CK},2\text{CK}}$ and $L \to \infty$, $|\phi_{\eta\eta}\rangle$ can be chosen by orthonormalizing $\{e_{i\eta}\}$. (ii) For each subspace $\mathcal{H}_i$, we straightforwardly obtain the optimal EW $X_i$ for $\mathcal{E}_F(\rho_i) = \text{Tr} X_i \rho_i$, where $\rho_i = I_i \rho I_i$ and $I_i$ is the identity operator of $\mathcal{H}_i$. For example, when $p_i \propto |\Psi_i\rangle \langle \Psi_i|$ and $|\Psi_i\rangle$ is the normalized Bell state, we choose $X_i = |2|\Psi_i\rangle \langle \Psi_i| - (2 - \log 2) I_i / \log 2$.
The universal scaling behavior of \( E \) is between the Kondo impurity and the bath electrons inside \( L \) at temperature \( T \) for single- (1CK) and two-channel Kondo (2CK) systems. (a,b) \( E \) versus \( T \) at \( L \to \infty \) for (a) 1CK and (b) 2CK. (c) \( E \) versus \( L \) at \( T = 0 \) for 1CK. Blue circles (green triangles) denote the lower (upper) bounds of \( E \). These bounds are close to each other, especially at \( T < T_{1CK,2CK} \) and \( L > \xi_{1CK} \), enough to predict scaling behavior. Insets: The universal scaling behavior of \( E \), versus \( T/T_{1CK} \) for (a) 1CK and (b) 2CK, and \( \xi_{1CK} \) or \( T/T_{2CK} \). The NRG parameters used for this plot and the expression of \( T_{1CK,2CK} \) are given in Ref. [35].

see Appendix. (iii) The sum \( X = \sum \lambda X_i \) of the two-qubit EWSs is our variational form [35]. \( X_i \) cannot detect off-diagonal blocks \( I_{\rho i} I_{\rho i} \), which are however made asymptotically small at \( T < T_{1CK} \) and \( L > \xi_{1CK} \) by appropriately choosing \( \langle \phi_{\nu} \rangle \).

Result. — We discuss the result of the temperature dependence of \( E \) at \( L \to \infty \) in Fig. 2. In both 1CK and 2CK, \( E \) shows maximal entanglement at \( T = 0 \), slowly decays with \( T < T_{1CK,2CK} \), and rapidly vanishes toward 0 at \( T \geq T_{1CK,2CK} \), exhibiting the crossover around \( T_{1CK,2CK} \). Particularly, at \( T \ll T_{1CK,2CK} \), the upper and lower bounds show the same universal power-law decay in each system,

\[
E_F \simeq 1 - a_1(T/T_{1CK})^2 \quad (1CK),
E_F \simeq 1 - a_2(T/T_{2CK})^2 \quad (2CK).
\]

2CK has more fragile macroscopic entanglement at scaled low temperature.

In Eq. (3), the scaling exponent is halved from 1CK to 2CK, reflecting different low-energy excitations between them. At \( T < T_{1CK,2CK} \), the thermal state \( \rho = \sum_i w_i |E_i \rangle \langle E_i| \) is governed by the \( |E_i\rangle \)s of \( E_i \simeq k_B T \), because of the competition between Boltzmann weight \( w_i \) and degeneracy. There are two factors suppressing \( E_F(\rho) \): (i) Each pure state \( |E_i\rangle \) is less entangled; \( E_F(|E_i\rangle) \simeq 1 - 2|S_{z,ii}|^2/\log 2 \) for \( E_i \ll k_B T_{1CK,2CK} \), where \( S_z \) is the impurity spin-\( z \) operator and \( S_{z,ii} = \langle E_i|S_z|E_i\rangle = (|\langle i|\rangle|^2 - |\langle \rangle|^2)/2 \). (ii) \( E_F \) satisfies the convexity, \( E_F(\sum_i w_i |E_i \rangle \langle E_i|) \leq \sum_i w_i E_F(|E_i\rangle) \). Using bosonization [34], we find that these two factors give the same exponent [35]. Here we explain the former factor. A pseudofermion operator \( c_{i}^\dagger \) describes the impurity spin as \( S_z = c_i^\dagger - 1/2 \). In 1CK, both \( c_i^\dagger \) and \( c_i \) couple to the bath. Since the coupling is energy dependent, each of \( c_i^\dagger \) and \( c_i \) gives a scaling factor \( \sim \sqrt{T/T_{1CK}} \), leading to \( 1 - E_F(|E_i\rangle) \sim |S_{z,ii}|^2 \sim (T/T_{1CK})^2 \). In 2CK, \( S_z \) is written as \( S_z = i \gamma_{-} \). Here, only a Majorana fermion \( \gamma_{-} \) couples to the bath, providing the factor \( \sim \sqrt{T/T_{2CK}} \), the other Majorana \( \gamma_{+} \) is decoupled and \( T \) independent. This causes the exponent halving in 2CK. It is non-Fermi liquid behavior.

The dependence of \( E_F \) on \( L \) is obtained for 1CK in Fig. 2(c). \( E_F(L \to \infty) - E_F(L) \) indicates the entanglement between \( x > L \) and the rest, which is lost in \( E_F(L) \); a proposal for Kondo cloud detection has similar motivation [24]. At \( T = 0 \), \( E_F = 1 \) at \( L \to \infty \) and decreases only slightly at \( L > \xi_{1CK} \), implying that Kondo cloud lies mostly (more than 90 %) inside \( \xi_{1CK} \). We find that the cloud has a long tail of the power law at \( L \gg \xi_{1CK} \),

\[
E_F \simeq 1 - b_1(\xi_{1CK}/L) \quad (1CK).
\]

We obtain the same power law [35] from Yosida's ground state [36].

At finite \( T \), the \( L \) dependence of \( E_F \) characterizes the thermal reduction of Kondo cloud. We find that the cloud size, within which the majority of the cloud lies, is \( \xi_{1CK} \) (almost insensitive to \( T \)) at \( T \lesssim T_{1CK} \), and decreases with \( T \) at \( T \gtrsim T_{1CK} \) [35]. Moreover, the two 1CK power-law decays in Eqs. (3) and (4) have different exponents, not connected by \( L \leftrightarrow \hbar v_F / 8 \) from the uncertainty principle, and they are additive at \( L \gg \xi_{1CK} \) and \( T \ll T_{1CK} \) as \( 1 - E_F \simeq a_1(T/T_{1CK})^2 + b_1(\xi_{1CK}/L) \). These unusual results imply that entanglement suppression by thermal effects has different mechanism from that by the partial trace over \( x > L \). The former reflects thermal entanglement suppression, while the latter measures the spatial extension of entanglement in thermal states.
Finally, in contrast to EoF, correlations between the impurity spin and a conduction electron spin at $L$ do not detect macroscopic entanglement, because of the entanglement monogamy that tracing out all bath electrons except the one at $L$ leaves only negligible entanglement. They measures the cloud tail differently from $E_{F}$; for example, the spin-spin correlation $\xi^{2}$ decays as $1/L^{2}$ at $L \gg \xi_{CK}$, and the concurrence does not detect Kondo cloud $\xi_{CK}$ Impurity entanglement entropy $[27, 28]$ detects macroscopic correlations, but it is not an entanglement measure; it decays as $T/T_{1CK}$, contrary to Appendix.— For any (normalized or unnormalized) state $\rho_{i}$ in $\mathcal{H}_{i}$, there exists the optimal witness operator $X_{i}$, which is dependent on $\rho_{i}$ and provides $E_{F}(\rho_{i}) = \text{Tr} \rho_{i}$, we derive $X_{i}$ $\xi_{F}(\rho_{i}) = f(x_{i})\text{Tr} \rho_{i}$. For the case of $\mathcal{C}(\rho_{i}) \neq 0$, the operator $X_{i}$ satisfies $\text{Tr} \rho_{i} = \mathcal{C}(\rho_{i}) = \sup \text{Tr} \mathcal{O}(2|\Psi_{i}| - I_{i})\mathcal{O}^{\dagger} \rho_{i}$, and is easily obtained $\mathcal{O}$ on $\mathcal{H}_{i}$ that makes $\text{Tr} \mathcal{O}(2|\Psi_{i}| - I_{i})\mathcal{O}^{\dagger} \rho_{i}$ the largest. Here $|\Psi_{i}\rangle = (|\eta\rangle|\phi_{\eta}^{+}\rangle + |\eta\rangle|\phi_{\eta}^{+}\rangle\sqrt{2} = \text{maximally entangled (}E_{F} = 1\text{)} Bell state. When $\mathcal{C}(\rho_{i}) = 0$, one chooses $X_{i} = x_{i} = 0$ (the null operator), since $E_{F}(\rho_{i}) = 0$. When $\rho_{i} \propto |\Psi_{i}\rangle_{\Psi_{i}}$, e.g., $|\Psi_{i}\rangle$ is the ground state in the case $T = 0$ and $L \rightarrow \infty$, we obtain $X_{i}$ $= \frac{2}{\log 2} |\Psi_{i}\rangle_{\Psi_{i}} - \left(\frac{2}{\log 2} - 1\right) I_{i}$, using $\mathcal{C}(\Psi_{i}) = 1$ and putting $\frac{df}{dx_{i}} x_{i} = 1/\log 2$ and $\mathcal{O} = I_{i}$ into Eq. (6) The sum $X = \sum_{i} X_{i}$ of the two-qubit witness operators is a witness operator for $\rho$, and provides a lower bound of $E_{F}(\rho) = \text{Tr} \left(\sum_{i} X_{i}\rho\right)$. A upper bound of $E_{F}(\rho)$ is obtained as $\sum_{j} \rho_{j} E_{F}(|\psi_{j}^{\dagger}\rangle)$, by finding $|\psi_{j}^{\dagger}\rangle$ which is similar to a state in $P_{X} = \sum_{i} X_{i}$, and satisfies $\sum_{j} \rho_{j} |\psi_{j}^{\dagger}\rangle|\psi_{j}^{\dagger}\rangle = \rho$, as mentioned before. To get better bounds, we optimize the choice of $|\phi_{\eta}\rangle$ and $|\eta\rangle$, based on the structure of $\rho$ $35$. We emphasize that the decomposition into the two-qubit subspaces allows us to avoid the impractical cost of computing EoF by Eq. (1); in virtue of the NRG scheme, the number of the two-qubit subspaces of only about $10^{4}$ is enough to cover a wide range of $T$ and $L$.

**Conclusion and perspective.**— We have proposed a viable approach for computing macroscopic entanglement in thermal mixed states. The macroscopic entanglement measured by EoF unveils the thermal decay and spatial profile of Kondo cloud, the essence of Kondo effects, and characterizes the non-Fermi liquid behavior of 2CK by Majorana fermions. Our study implies that EoF is a good tool for quantifying macroscopic quantumness in many-body mixed states; the original operational meaning of EoF is a nonregularized entanglement cost in quantum information.

Our results indicate that the macroscopic entanglement characterizes the new aspects of many-body systems at finite temperature, which cannot be accessed by conventional means and by entanglement entropy. For example, it can identify the spatial extension of quantum correlations, the competition between the coexisting quantum and classical correlations induced by thermal effects or environments, and the fate of the zero-temperature correlations (e.g., topological order and quantum criticality) at finite temperature.

Our approach is optimized for computing entanglement between a few impurities and a macroscopic sub-system, and directly applicable to various quantum impurity problems. It is desirable to further develop our approach to study entanglement between macroscopic subsystems, using appropriate EWs. Moreover, our approach is in principle applicable to all types of convex-roof entanglement $[13, 15]$, including multipartite entanglement $[11, 12]$, and useful for experimental entanglement detection $[13, 31]$.

---

[1] L. Amico, R. Fazio, A. Osterloh, and V. Vedral, Rev. Mod. Phys. 80, 517 (2008).
[2] V. Vedral, Nature 453, 1004 (2008).
[3] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. 81, 865 (2009).
[4] J. Eisert, M. Cramer, and M. B. Plenio, Rev. Mod. Phys. 82, 277 (2010).
[5] A. Kitaev and J. Preskill, Phys. Rev. Lett. 96, 110404 (2006).
[6] M. Levin and X.-G. Wen, Phys. Rev. Lett. 96, 110405 (2006).
[7] G. Vidal, J. I. Latorre, E. Rico, and A. Kitaev, Phys. Rev. Lett. 90, 227902 (2003).
[8] P. Calabrese and J. Cardy, J. Phys. A: Math. Theor. 42, 504005 (2009).
[9] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Weber, Phys. Rev. Lett. 87, 167902 (2001).
K. Wootters, Phys. Rev. A 54, 3824 (1996).
[10] B. Röthlisberger, J. Lehmann, and D. Loss, Phys. Rev. A 80, 042301 (2009).
[11] S.-S. B. Lee and H.-S. Sim, Phys. Rev. A 85, 022325 (2012).
[12] S. Ryu, S.-S. B. Lee, and H.-S. Sim, Phys. Rev. A 86, 042324 (2012).
[13] O. Gühne and G. Tóth, Phys. Rep. 474, 1 (2009).
[14] M. B. Plenio, Science 324, 342 (2009).
[15] M. B. Plenio and S. Virmani, Quant. Inf. Comp. 7, 1 (2007).
[16] G. Vidal and R. F. Werner, Phys. Rev. A 65, 032314 (2002).
[17] M. B. Plenio, Phys. Rev. Lett. 95, 090503 (2005).
[18] A. Bayat, P. Sodano, and S. Bose, Phys. Rev. B 81, 064429 (2010).
[19] A. Bayat, S. Bose, P. Sodano, and H. Johannesson, Phys. Rev. Lett. 109, 066403 (2012).
[20] P. Calabrese, J. Cardy, and E. Tonni, Phys. Rev. Lett. 109, 130502 (2012).
[21] A. C. Hewson, The Kondo Problems to Heavy Fermions (Cambridge University Press, Cambridge, 1993).
[22] I. Affleck, Perspectives of Mesoscopic Physics (World Scientific, 2010), pp. 1-44.
[23] A. K. Mitchell, M. Becker, and R. Bulla, Phys. Rev. B 84, 115120 (2011).
[24] J. Park, S.-S. B. Lee, Y. Oreg, and H.-S. Sim, Phys. Rev. Lett. 110, 246603 (2013).
[25] L. Borda, Phys. Rev. B 75, 041307(R) (2007).
[26] A. Holzner, I. P. McCulloch, U. Schollwöck, J. von Delft, and F. Heidrich-Meisner, Phys. Rev. B 80, 205114 (2009).
[27] E. S. Sørensen, M.-S. Chang, N. Laflorencie, and I. Affleck, J. Stat. Mech. P08003 (2007).
[28] E. Eriksson and H. Johannesson, Phys. Rev. B 84, 041107(R) (2011).
[29] F. G. S. L. Brandão, Phys. Rev. A 72, 022310 (2005).
[30] J. Eisert, F. G. S. L. Brandão, and K. M. R. Audenaert, New J. Phys. 9, 46 (2007).
[31] H. S. Park, S.-S. B. Lee, H. Kim, S.-K. Choi, and H.-S. Sim, Phys. Rev. Lett. 105, 230404 (2010).
[32] A. Weichselbaum and J. von Delft, Phys. Rev. Lett. 99, 076402 (2007).
[33] R. Bulla, T. A. Costi, and T. Pruschke, Rev. Mod. Phys. 80, 395 (2008).
[34] G. Zaránd and J. von Delft, Phys. Rev. B 61, 6918 (2000).
[35] See Supplemental Material for the details of our approach and some supplementary results.
[36] K. Yosida, Phys. Rev. 147, 223 (1966).
[37] S. Oh and J. Kim, Phys. Rev. B. 73, 052407 (2006).
[38] W. K. Wootters, Phys. Rev. Lett. 80, 2245 (1998).
Here we provide the details of our approaches and some supplementary results. In Sec. \[A\] we briefly introduce the Kondo models. In Sec. \[B\] we describe how to numerically construct the thermal mixed states of the Kondo models by NRG, and give the NRG parameters. In Sec. \[C\] we describe in details the way of tracing out the subsystem in \(x > L\), within the NRG formalism. In Sec. \[D\] we derive the “two-qubit” EW \(X_i\) given in Eq. (5) in the main text. In Sec. \[E\] we prove that the sum \(X = \sum_i X_i\) is a valid EW, and discuss how to choose the bath basis \(\{\phi_{i\eta}\}_{i\eta}\). In Sec. \[F\] we give the way to obtain the upper bound of \(E_F(\rho)\). In Sec. \[G\] we analyze the scaling behavior of the thermal suppression of \(E_F\) in Eq. (3) in the main text, using the finite-size bosonization method. In Sec. \[H\] we reproduce the long-tail scaling of \(E_F(L)\) in Eq. (4) in the main text, using the Yosida’s variational ground state. In Sec. \[I\] we give the computation result of EO\(F\) for 1CK when both \(T\) and \(L\) are finite, to discuss the size of the Kondo cloud at finite \(T\). We also address that, at \(T \ll T_{1\text{CK}}\) and \(L \gg \xi_{1\text{CK}}\), two power-law decays are additive.

### A. Kondo Hamiltonian

In 1CK (2CK), a spin-1/2 impurity is antiferromagnetically coupled with the spin(s) of a single channel (two channels) of the conduction electron bath at the impurity site \([21]\). Without loss of generality, we consider a semi-infinite one-dimensional bath ranging from \(x = 0\) (the impurity site) to \(x \to \infty\). Its Hamiltonian is

\[
H = J \sum_{\alpha} \vec{S} \cdot \vec{s}_{\alpha} + \sum_{\alpha \sigma} \epsilon_{\alpha} c_{\alpha \sigma}^\dagger c_{\alpha \sigma},
\]

where \(J\) is the coupling strength, \(\vec{S}\) is the impurity spin operator, \(\vec{s}_{\alpha} = \sum_{k \sigma \sigma'} c_{\alpha \sigma}^\dagger \sigma \sigma' c_{\alpha \sigma'} / 2\) is the electron spin operator in channel \(\alpha \in \{1, M\}\) at the impurity site \((x = 0)\), \(M = 1\) (2) for 1CK (2CK), \(\sigma\) is Pauli matrix, \(c_{\alpha \sigma}^\dagger\) creates an electron with spin \(\sigma\), momentum \(k\), and energy \(\epsilon_{\alpha} = \hbar v_F (k - k_F) \in (-D, D)\) in \(\alpha\), constant density of states \(\nu = 1/2D\), \(k_F\) is Fermi momentum, and \(D\) is the bandwidth. In this work, we consider the following case: The two channels of 2CK have the same coupling strength, and there is no external magnetic field. We use 1CK Kondo temperature \(k_B T_{1\text{CK}} = D \sqrt{\nu J} e^{-1/\nu J}\), and determine \(T_{2\text{CK}}\) from energy-eigenvalue convergence in NRG.

### B. Density matrix by NRG

In NRG \([33]\), each channel is logarithmically discretized and mapped onto a semi-infinite tight-binding chain of length \(N\), whose Hamiltonian is \(H_N = J \vec{S} \cdot \sum_{\alpha} \vec{s}_{\alpha} + \sum_{\alpha \sigma} \epsilon_{\alpha} f_{\alpha \sigma}^\dagger f_{\alpha \sigma} + \sum_{\alpha n} (t_n f_{\alpha n}^\dagger f_{\alpha n+1} + \text{H.c.})\), where \(f_{\alpha n}^\dagger\) creates an electron in the single-particle state \(|\alpha n\sigma\rangle\) of spin \(\sigma\) in channel \(\alpha\) at site \(n\), \(t_n \sim DA^{-n/2}\), and \(A\) is the discretization parameter. \(|\alpha n\sigma\rangle\) has energy \(\sim DA^{-n/2}\) and extends over length \(\sim k_F^{-1} A^{n/2}\). \(H_N\) is iteratively diagonalized, based on the energy-scale hierarchy. At each \((n\text{-th})\) iteration step, only the lowest-lying energy eigenstates \(\{|E_{ni}\}_i\) of the Hamiltonian of the step are kept to construct the next-step Hamiltonian, while the rest \(\{|E_{ni}\}_i\) is discarded. The discarded states are the energy eigenstates of \(H_N\), \(H_N|E_{ni}\rangle \otimes \otimes_{\alpha' n' > n} |s_{\alpha n'}\rangle \approx E_{ni}^D |E_{ni}\rangle \otimes \otimes_{\alpha' n' > n} |s_{\alpha n'}\rangle\), where \(|s_{\alpha n} = 0, \uparrow, \downarrow, \uparrow\rangle\) denote the occupation basis states of site \(n\) and channel \(\alpha\).

The equilibrium state at temperature \(T\) is constructed \([32]\) as

\[
\rho(T) = \sum_n \rho_n \otimes I_{>n}, \quad \rho_n = \sum_i e^{-E_{ni}/k_B T} |E_{ni}\rangle \langle E_{ni}^D|, \quad I_{>n} = \bigotimes_{\alpha' n' > n} \sum_s |s_{\alpha n'}\rangle \langle s_{\alpha n'}|,
\]

where \(k_B\) is Boltzmann constant and \(Z\) is the partition function. Each block \(\rho_n\) covers energy \(\sim DA^{-n/2}\) and length \(\sim k_F^{-1} A^{n/2}\). \(\text{Tr}(\rho_n \otimes I_{>n})\) is maximal near \(n = n_T \equiv -2 \log_A (k_B T / D)\) due to the competition between Boltzmann factor \(e^{-DA^{-n/2}/k_B T}\) and degeneracy \(\text{Tr} I_{>n} = 4^{N-n}\). In this work, we choose \(A = 4, J/D = 0.3\), and the number of kept states \(\lesssim 300\) at each iteration, and use the \(z\)-averaging \([33]\) with \(z = 0\) and 0.5.
FIG. S3. $|P_L|_{nn'}$ (upper panels) and $p_n \equiv |P_L|_{nn}$ (lower panels) for (a) $\Lambda = 4$ and (b) $\Lambda = 8$. $L$ is chosen to be $\Lambda$ dependent as $L = k_F^{-1}A^*$ for comparison; $n_L = 16$ for both of (a) and (b). Note that for $\Lambda = 4$, the largest off-diagonal element of $|P_L|_{nn'}$ is $\sim 0.07$ and the other off-diagonal ones are smaller by one or more orders. We choose $z = 0$.

C. Partial trace over $x > L$

We develop a way of obtaining, from NRG state $\rho(T)$, the reduced density matrix $\rho(T, L)$ of the impurity and electrons in $x \leq L$, by tracing out states in $x > L$. We use the projector to $x \leq L$,

$$P_L = \frac{1}{a} \int_0^L \text{d}x |x\rangle \langle x|.$$  \hspace{1cm} (S9)

$|x\rangle$ is the state spatially localized at $x$ and $a$ is lattice constant. One has $\langle an\sigma|P_L|\alpha'n'\sigma'\rangle = \delta_{\alpha n}\delta_{\sigma \sigma'}|P_L|_{nn'}$. The matrix $|P_L|_{nn'}$ is real symmetric and almost diagonal (see Fig. S3). Its diagonal part $|P_L|_{nn}$ is finite for $n \lesssim n_L \equiv 2\log k_F L$, and vanishes for $n \gtrsim n_L$, reflecting the length scales of NRG sites. Off-diagonal parts $|P_L|_{nn'}$ are much smaller than diagonal ones, and decrease with increasing $\Lambda$, since spatial separation between $(n\sigma)$'s increases. Based on this observation, we neglect the off-diagonal elements. The insensitivity of our computation of $E_F(\rho)$ to $\Lambda$ implies that this is a good approximation.

Neglecting the off-diagonal parts of $|P_L|_{nn'}$, we decompose $|an\sigma\rangle$ into the states of $x \leq L$ and $x > L$ as $f_{an\sigma}^\dagger = \sqrt{p_n}f_{an\sigma,\text{in}}^\dagger + \sqrt{1-p_n}f_{an\sigma,\text{out}}^\dagger$, where $p_n \equiv |P_L|_{nn}$. Then the many-body occupation basis state $|s_{an}\rangle$ of site $n$ and channel $\alpha$ is expressed as

$$|0_{an}\rangle = |0_{an}\rangle^\text{in},$$  

$$|\uparrow_{an}\rangle = f_{an\uparrow}^\dagger|0_{an}\rangle,$$

$$|\downarrow_{an\uparrow}\rangle = \sqrt{p_n}|0_{an}\rangle + \sqrt{1-p_n}f_{an\uparrow,\text{out}}^\dagger|0_{an}\rangle,$$

$$|\downarrow_{an\downarrow}\rangle = f_{an\downarrow}^\dagger|0_{an}\rangle,$$

$$|\uparrow_{an\downarrow}\rangle = \sqrt{p_n}|0_{an}\rangle + \sqrt{1-p_n}f_{an\downarrow,\text{out}}^\dagger|0_{an}\rangle,$$

$$|\uparrow_{an}\rangle = f_{an\uparrow}^\dagger f_{an\downarrow}^\dagger|0_{an}\rangle,$$

where $|s_{an}\rangle^\text{in(out)}$ describes the occupation $s = 0, \uparrow, \downarrow, \uparrow$ in the $x \leq L (x > L)$ part of site $n$ and channel $\alpha$. The Hilbert space of $x \leq L (x > L)$ is spanned by $\bigotimes_{\alpha,n} \{|s_{an}\rangle^\text{in(out)}\}_s$. In this representation, tracing out $x > L$ is equivalent to
partial trace over $\otimes_{\alpha,n}\{|s_{\alpha,n}^{\text{out}}\rangle\rangle$.

The partial trace over $\otimes_{\alpha,n}\{|s_{\alpha,n}^{\text{out}}\rangle\rangle$ can be efficiently done for $\rho(T)$ in Eq. (88) as $\rho(T, L) = T_{\alpha=1}^{\text{out}} \cdots T_{\alpha=N}^{\text{out}} \rho(T)$, where $T_{\alpha=1}^{\text{out}}(\cdot) \equiv T_{\alpha=1}^{\text{out}}(\cdot)$ is the partial trace applied to site $n$ and $T_{\alpha=n}^{\text{out}}(\cdot) = \sum_s |s_{\alpha,n}^{\text{out}}\rangle \langle s_{\alpha,n}^{\text{out}}|$ . By choosing appropriate basis states $|r_n\rangle \in \text{span} \otimes_{\alpha,n' \leq n} \{|s_{\alpha,n'}^{\text{in}}\rangle\rangle$, we express $\rho(T, L)$, the result of the partial trace, as the block diagonal form of

$$
\rho(T, L) = \sum_n \rho_n^{\text{in}} \otimes I_{\geq n}^{\text{in}}, \quad \rho_n^{\text{in}} = \sum_{r_n|m} r_{n|m} \langle r_n| \langle r_n|, \quad I_{\geq n}^{\text{in}} = \bigotimes_{\alpha,n' > n} R_{\alpha,n'} \langle s_{\alpha,n'}^{\text{in}}| \langle s_{\alpha,n'}^{\text{in}}|
$$

This form corresponds to Eq. (88) with $\rho_n \rightarrow \rho_n^{\text{in}}, I_{\geq n} \rightarrow I_{\geq n}^{\text{in}}$, and $|E_{\alpha,i}^{\text{in}}\rangle \otimes \otimes_{\alpha,n' > n} |s_{\alpha,n'}\rangle \rightarrow |r_n\rangle \otimes \otimes_{\alpha,n' > n} |s_{\alpha,n'}^{\text{in}}\rangle$. $\text{Tr}(\rho_n^{\text{in}} \otimes I_{\geq n}^{\text{in}})$ is maximal near $n = \min\{n_1, n_L\}$. As $\rho_n$’s are defined not spatially but energetically, $\rho_n^{\text{in}}$ can be contributed from many $\rho_n$’s. In this form, we generalize the concept of the kept and discarded states used for $\rho(T)$ into $\rho(T, L)$. In each step of constructing $\rho_n^{\text{in}}$, $|r_n\rangle$’s hold for all convex-roof entanglement measures and beyond two qubits [11] and it is consistent with Eq. (2). This is useful, as the “two-qubit” state $\rho_i = I_i \rho f_i$ in the main text is usually unnormalized.

The derivation of $X_i$ starts with the optimal witness operator $X_i^C$ for $C(\rho_i)$. In the case of $C(\rho_i) \neq 0$, it is obtained [11] [31] as

$$
C(\rho_i) = \text{Tr} X_i^C \rho_i = \sup_{\mathcal{O}} \text{Tr}[\mathcal{O}(2|\Psi_i\rangle \langle \Psi_i| - I_i)\mathcal{O}^\dagger \rho_i], \quad \mathcal{O} = \mathcal{O}_\eta \otimes \mathcal{O}_\phi, \quad (S10)
$$

where $O_\eta$ and $O_\phi$ are local operators with determinant 1, acting on $\{|\eta\rangle, |\eta\rangle\rangle$ and $\{|\phi_1\rangle, |\phi_1\rangle\rangle$}, respectively. Here $|\Psi_i\rangle = (|\eta\rangle \langle \eta| + |\psi\rangle \langle \psi|)/\sqrt{2}$ is maximally entangled ($\mathcal{E}_\mathcal{F} = 1$) Bell state. In Eq. (S10), $X_i^C$ is found by searching the optimal SLOCC (stochastic local quantum operations and classical communications in quantum information theory) operator $\mathcal{O}$ on $\mathcal{H}_i$ that makes $\text{Tr}[\mathcal{O}(2|\Psi_i\rangle \langle \Psi_i| - I_i)\mathcal{O}^\dagger \rho_i]$ the largest. The form of Eq. (S10) captures the invariance of $C$ under SLOCC. In the case of $C(\rho_i) = 0$, on the other hand, we choose $X_i^C = 0$ (the null operator). $X_i^C$ provides $C(\rho_i) = \text{Tr} X_i^C \rho_i$. Searching the optimal operator $\mathcal{O}$ can be easily done by the singular value decomposition of the local operators as $\mathcal{O}_{\eta(\phi)} = U_{\eta(\phi)} F_{\eta(\phi)} U_{\eta(\phi)}$, where $U_{1,2}$ are $2 \times 2$ local unitary operators and $F$ is a local filtering operator [31]; in the matrix representation, $F$ is written as $(f_{ij})$ with real $f$.

$X_i$ is obtained from $X_i^C$ and $\mathcal{E}_\mathcal{F}(\rho_i/\mathcal{Tr}\rho_i) = f(C(\rho_i/\mathcal{Tr}\rho_i))$. As $f(x)$ is monotonically increasing and convex, one has $f(x) \geq f(x_0) + \frac{df}{dx}\bigg|_{x=x_0} (x-x_0)$ at any $x_0$ and $x$. Substituting $x_0 = C(\rho_i/\mathcal{Tr}\rho_i)$, $f(x) \rightarrow X_i$, $x \rightarrow X_i^C$, and $1 \rightarrow I_i$, we choose

$$
X_i = f(x_0) I_i + \frac{df}{dx}\bigg|_{x=x_0} (X_i^C - x_0 I_i), \quad x_0 = C(\rho_i/\mathcal{Tr}\rho_i) = C(\rho_i)/\mathcal{Tr}\rho_i. \quad (S11)
$$

This operator $X_i$ is a witness operator for $\mathcal{E}_\mathcal{F}$, for any state $\rho' \in \mathcal{H}_i$, $\text{Tr} X_i \rho' \leq \mathcal{E}_\mathcal{F}(\rho')$; one can check $\text{Tr} X_i \rho' \leq \text{Tr} f(x_0)/\mathcal{Tr}\rho' + \frac{df}{dx}\bigg|_{x=x_0} (C(\rho') - x_0 \text{Tr}\rho') \leq f(C(\rho')/\mathcal{Tr}\rho')/\mathcal{Tr}\rho' = \mathcal{E}_\mathcal{F}(\rho'/\mathcal{Tr}\rho')/\mathcal{Tr}\rho' = \mathcal{E}_\mathcal{F}(\rho')$, using the convexity of $f(x)$ and Eq. (2). Moreover, it is easy to show that $X_i$ satisfies $\text{Tr}(X_i \rho_i) = \mathcal{E}_\mathcal{F}(\rho_i)$, using $\text{Tr}(X_i^C \rho_i) = C(\rho_i)$. Therefore $X_i$ is the optimal witness operator for $\mathcal{E}_\mathcal{F}(\rho_i)$. This proves Eq. (5).

E. Witness operator $X = \sum_i X_i$

In the main text, we divide the whole Hilbert space $\mathcal{H}$ into “two-qubit” subspaces $\mathcal{H}_i \equiv \text{span}\{|\eta\rangle \otimes |\phi_{n'}\rangle\rangle$ and obtain the optimal witness operator $X_i$ for $\mathcal{E}_\mathcal{F}(\rho_i)$, directly from $X_i; \rho_i = I_i \rho f_i$ is the projection of $\rho$ to $\mathcal{H}_i$ and
\(\Tr X_i \rho_i = \mathcal{E}_F(\rho_i)\). We here (i) prove that \(X = \sum_i X_i\) is a witness operator for \(\mathcal{E}_F(\rho)\), namely that \(\Tr X \rho\) is a lower bound of \(\mathcal{E}_F(\rho)\), and also (ii) discuss a strategy how to optimize \(X_i\)’s.

The task (i) is equivalent, according to Eq. (2), to proving that \(\Tr (X|\psi\rangle \langle \psi|) \leq \mathcal{E}_F(|\psi\rangle)\) for any normalized pure state \(|\psi\rangle\) in \(\mathcal{H}\). To prove it, we decompose \(|\psi\rangle = \sum_i |\psi_i\rangle\), where \(|\psi_i\rangle\) is the projection of \(|\psi\rangle\) onto \(\mathcal{H}_i\). Applying Schmidt decomposition to \(|\psi_i\rangle\), \(|\psi_i\rangle = c_{\psi}(|\phi_{\psi}\rangle + c_{\psi}|\phi_{\psi}'\rangle\) and \(|\phi_{\psi}\rangle\) = \(\sum_i |c_{\psi}^i|^2\) \(\delta_{\psi\psi'}\) has one has \(\mathcal{E}_F(|\psi\rangle) = \langle \sum_i |c_{\psi}^i|^2 \rangle = \sum_i (|c_{\psi}^i|^2)^2 \) \(\delta_{\psi\psi'}\) and \(\rho(x) = -x \log_2 x - (1 - x) \log_2 (1 - x)\). It satisfies \(\mathcal{E}_F(|\psi\rangle) \geq \sum_i \mathcal{E}_F(|\psi_i\rangle) \sum_i |\psi_i| X_i |\psi_i\rangle = \Tr(\sum_i X_i |\psi\rangle \langle \psi|)\); the first inequality is from the concavity of \(h(x)\), \(h(|\sum c_{\psi}^i|^2) \geq \sum_i (|c_{\psi}^i|^2)^2 h((|c_{\psi}^i|^2)^2)\)(\(\bar{c}_{\psi}+c_{\psi}^{\dagger}\)), and also from \(\mathcal{E}_F(|\psi_i\rangle) = \langle \hat{\phi}|\psi_i\rangle \mathcal{E}_F(|\psi_i\rangle) / \sqrt{\langle \hat{\phi}|\psi_i\rangle}\), and the second from the fact that \(X_i\) is a witness operator for the EoP of two-qubit states (which was proved above). This proves that \(X\) is a witness operator for \(\mathcal{E}_F(\rho)\).

Next, we discuss a strategy how to find \(X = \sum_i X_i\) that provides a better lower bound of \(\mathcal{E}_F(\rho)\). One needs to first decompose \(\mathcal{H}\) into \(\mathcal{H}_i\)’s. Among many possible ways for it, we choose a NRG-based way. In this way, we decompose \(\rho\) into “units”, and choose the basis state span\{|\eta\| \otimes |\phi_{ii'}\rangle\}_{i\neq i'} of each unit. Each unit has one or a few successive NRG diagonal blocks of \(\rho\), and different units have no overlap; the number of blocks in a unit is chosen to have a lower better bound of \(\mathcal{E}_F(\rho)\). Then \{|\phi_{ii'}\rangle, |\phi_{ii'}\rangle\} constitutes \(\mathcal{H}_i\). This way is naturally expected to lead to a good lower bound, as the NRG blocks capture the main physics. After choosing \(\mathcal{H}_i\)’s, we find the optimal witness \(X_i\) (equivalently \(X_i^T\)) for \(\mathcal{E}_F(\rho_i) = \Tr X_i \rho_i\), following Eqs. (S10) and (S11). We skip other technical details of finding \(X\), such as how to choose the basis states \{|\phi_{ii'}\rangle\}_{i\neq i'}.

\[\mathbf{F. \ Upper \ bound \ of \ \mathcal{E}_F(\rho)}\]

In the above, we find \(X = \sum_i X_i\) that provides the best lower bound of \(\mathcal{E}_F(\rho)\) within the form utilizing Eq. (S11). We call this operator as \(X_{\rho^*}\). A good lower bound is also obtained from \(X_{\rho^*}\), by finding a set of pure states \(\mathbb{P}_{X_{\rho^*}} = \{|\psi\rangle \langle \psi| X_{\rho^*} |\psi\rangle = \mathcal{E}_F(|\psi\rangle)\}\) and a decomposition \(\rho = \sum_i \rho_i |\psi_i\rangle \langle \psi_i|\) where each \(|\psi_i\rangle\) is sufficiently similar to an element of \(\mathbb{P}_{X_{\rho^*}}\). We suggest below a systematic way of finding the decomposition \(\rho = \sum_i \rho_i |\psi_i\rangle \langle \psi_i|\).

To find the decomposition, we diagonalize \(\rho = \sum_i \rho_i |\varphi_i\rangle \langle \varphi_i|\) is generated by a left-unitary matrix \(U \sqrt{\mathbb{Q}} |\varphi_i\rangle \langle \varphi_i| U\), \(\sum_i \mathbb{U}_{il} \sqrt{\mathbb{Q}} |\varphi_i\rangle \langle \varphi_i| U = I\). To generate \(|\varphi_i\rangle\) close to \(\mathbb{P}_{X_{\rho^*}}^T = \{|\psi_i\rangle\}\), we introduce a matrix \(W_i\), \(W_i|\psi\rangle = \frac{\sqrt{\mathbb{P}_{X_{\rho^*}}}}{U} |\psi\rangle\), and obtain its singular value decomposition of \(W = V_i \Sigma V_i^\dagger\), where \(y_i\) and \(y_2\) are the variables to be optimized. Here, \(p_i^2\)’s are chosen to satisfy \(\rho \sim \sum_i p_i |\psi_i\rangle \langle \psi_i|\). Then, we choose \(U\) as \(U = V_i^\dagger \Sigma^{-1}\) and use it to obtain \(|\psi^T_i\rangle\) via \(\sqrt{\mathbb{P}_{X_{\rho^*}}} = \sum_i U_i \sqrt{\mathbb{P}_{X_{\rho^*}}} |\psi_i\rangle\).

Finally, we optimize \(y_1\) and \(y_2\) to minimize \(\sum_i p_i^2 \mathcal{E}_F(|\psi^T_i\rangle)\). The minimum value of \(\sum_i p_i^2 \mathcal{E}_F(|\psi^T_i\rangle)\) is a good upper bound of \(\mathcal{E}_F(\rho)\).

In the above way of finding a upper bound, it takes heavy numerical cost to handle \(\rho\) as a whole, since \(\rho\) has a large size. To avoid the heavy cost, we decompose \(\rho = \sum_n \rho_n\) into the NRG blocks \(\rho_n\)’s (or the units of a few successive blocks), construct a witness operator \(X_n\) for \(\mathcal{E}_F(\rho_n)\), and find a good upper bound \(\mathcal{E}_n\) of \(\mathcal{E}_F(\rho_n)\), using \(\mathbb{P}_{X_n}\) as mentioned above. The sum \(\sum_n \mathcal{E}_n\) of the upper bound of \(\mathcal{E}_F(\rho_n)\) over \(n\)’s provides a good upper bound of \(\mathcal{E}_F(\rho)\). Note that \(\sum_n X_n\) is not necessarily a witness operator of \(\rho\); it is because \(X_n\) is not necessarily constructed by basis states orthogonal between different NRG blocks (or units), contrary to \(X = \sum_i X_i\).

\[\mathbf{G. \ T \ dependence \ of \ \mathcal{E}_F \ from \ bosonization}\]

We here confirm the universal power-law thermal decay of \(\mathcal{E}_F\), using finite-size bosonization and refermionization methods [51], and attribute the power-law exponents different between 1CK and 2CK to Majorana fermions emerging in 2CK.

For 1CK and 2CK, the thermal state has the form of \(\rho = \sum_i w_i |E_i\rangle \langle E_i|\); \(w_i\) is Boltzmann weight. \(|E_i\rangle = b_{i\uparrow} |\uparrow\rangle \langle \downarrow| e_{i\uparrow}\rangle + b_{i\downarrow} |\downarrow\rangle \langle \uparrow| e_{i\downarrow}\rangle\) is an energy eigenstate with energy \(E_i\) and an eigenstate of the total (impurity and bath) spin–z operator simultaneously. Bath states \(|\phi_{i\uparrow}\rangle\) satisfy \(|\phi_{i\uparrow}\rangle |\phi_{i\downarrow}\rangle = 0\) because \(|\phi_{i\uparrow}\rangle\) and \(|\phi_{i\downarrow}\rangle\) have different spin–z quantum numbers, while \(|\phi_{i\uparrow}\rangle |\phi_{i\downarrow}\rangle \neq 0\) in general. We focus on \(|E_i\rangle\)’s with \(E_i \sim k_B T\), as they govern the properties of \(\rho\); this is due to the competition between degeneracy and Boltzmann weight. Using the bosonization, we will later show that for \(E_i, E_i \sim k_B T \ll k_B T_{1CK,2CK}, S_{z,iv} \equiv \langle E_i | S_z | E_i \rangle\) and \(S_{-z,iv} \equiv \langle E_i | -S_z | E_i \rangle\) satisfy

\[S_{z,iv}, S_{-z,iv} \sim \begin{cases} T/T_{1CK}, & \text{for 1CK,} \\ \sqrt{T/T_{2CK}}, & \text{for 2CK.} \end{cases}\]

(S12)
The $S_z$ and $S_-$ impurity spin operator, $S_z \equiv (|\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|)/2$ and $S_- \equiv |\downarrow\rangle\langle\uparrow|$, have entanglement information. $S_{z,ii}$ connects with $\tilde{E}_F(|E_{i1}\rangle)$. $|E_{i1}\rangle$ is maximally entangled when $S_{z,ii} = (|\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|)/2 = 0$, while it is separable when $S_{z,ii} = \pm 1/2$. We find that for $E_{i1}/k_B T \ll T_{1\text{CK}}, 2\text{CK}$, $|S_{z,ii}\rangle \ll 1/2$ and $\tilde{E}_F(|E_{i1}\rangle) \approx h (\frac{1}{2} + S_{z,ii}) \approx 1 - 2|S_{z,ii}|^2/\log 2$, where $h(x) = -x \log_2 x - (1 - x) \log_2 (1 - x)$. On the other hand, $S_{z,ii} = 0$ and $S_- = 0$ have the information of state overlap $\langle e_{ii'}|e_{ii''}\rangle$. From Eq. (S12) and $\langle E_{i1}|E_{i'}\rangle = \delta_{ii'}$, we find

$$\langle e_{ii'}|e_{ii''}\rangle - \delta_{ii'}\delta_{ii''} \propto \begin{cases} \frac{T}{T_{1\text{CK}}} \sqrt{\frac{T}{2T_{2\text{CK}}}}, & \text{for 1CK,} \\ \frac{1}{2} \log 2 & \text{for 2CK.} \end{cases}$$ (S13)

The overlap results in entanglement reduction in a pure-state mixture. $\tilde{E}_F(\sum_i w_i|E_{i1}\rangle\langle E_{i1}|) \leq \sum_i w_i \tilde{E}_F(|E_{i1}\rangle)$. From Eq. (1), $\tilde{E}_{F,0} = \sum_i w_i \tilde{E}_F(|E_{i1}\rangle) \geq \sum_i w_i (1 - 2|S_{z,ii}|^2/\log 2)$ is a upper bound of $\tilde{E}_F(\rho)$. The upper bound $\tilde{E}_{F,0}$ and Eq. (S12) agree with the power law in Eq. (3).

We also confirm Eq. (3) using a lower bound of $\tilde{E}_F(\rho)$. We consider a witness operator $X'$, $X' = \sum_i \left[ \frac{2}{\log 2} |\Psi_i\rangle\langle\Psi_i| - \left( \frac{2}{\log 2} - 1 \right) I_i \right]$, $|\Psi_i\rangle = \frac{1}{\sqrt{2}} (|\uparrow\rangle|\phi_{ii}\rangle + |\downarrow\rangle|\phi_{ii}\rangle)$, $I_i = \sum_{\eta\gamma} |\eta\rangle\langle\eta| \otimes |\phi_{ii'}\rangle\langle\phi_{ii'}|$. This has the same form as Eq. (6). Here, $|\phi_{ii}\rangle$'s are the orthonormal states obtained by applying the Gram-Schmidt orthonormalization process to the states $\{|e_{ii}\rangle\}$. Because $\langle e_{ii'}|e_{ii''}\rangle = \delta_{ii'}\delta_{ii''}$ is very small at $T \ll T_{1\text{CK}}, 2\text{CK}$ as in Eq. (S13), $|\phi_{ii}\rangle$ little deviates from $|e_{ii}\rangle$ as $b_{ii'}|e_{ii}\rangle = (|\phi_{ii}\rangle + |\delta_{ii'}\rangle)/\sqrt{2}$. The expectation value $\text{Tr} X' \rho$ is a lower bound of $\tilde{E}_F(\rho)$. After some computation, we find

$$\text{Tr} X' \rho = 1 - \frac{1}{\log 2} \sum_{ii'} w_i (|\phi_{ii'}\rangle\langle\phi_{ii'}| + |\phi_{ii'}\rangle\langle\phi_{ii'}|)^2$$

Applying $|\delta_{ii}\rangle \propto T/T_{1\text{CK}}$ for 1CK and $|\delta_{ii}\rangle \propto \sqrt{\frac{T}{T_{2\text{CK}}}}$ for 2CK in Eq. (S13), we find that $\text{Tr} X' \rho$ satisfies Eq. (3). This analytic derivation of the same universal power-law behavior of the upper and lower bounds $\tilde{E}_{F,0}$ and $\text{Tr} X' \rho$ strongly supports our numerical result of Eq. (3).

For a complete proof, we now derive Eq. (S12) and discuss the difference between 1CK and 2CK. We first consider 2CK. In 2CK, the electron bath has the four degrees of freedom, total charge, total spin, charge difference between the channels, and spin difference between the channels. According to the bosonization and reformation along Emery-Kivelson line [34], the degree of freedom from the spin difference (decoupled from the others) is described by the resonant-level model, $H_x^{2\text{CK}} = c_d^\dagger c_d + \sum_k \epsilon_k c_k^\dagger c_k + \sqrt{\Delta T} \sum_k (c_k^\dagger + c_k)(c_d - c_d^\dagger)$, where $c_d^\dagger$ creates a pseudofermion in the resonant level coupled to a reservoir of pseudofermions (with momentum $k$ and energy $\epsilon_k$) created by $c_k^\dagger$, $\Delta$ is the level spacing of the resonance, and $\epsilon$ means normal ordering. $\Gamma$ is the broadening of the resonance and plays the role of Kondo temperature, $\Gamma = k_B T_{2\text{CK}}$. We choose $\Delta$ as $\Delta = \delta - k_B T$ to focus on energy scale $\sim k_B T$. $c_d^\dagger (c_d)$ corresponds to impurity spin raising operator $S_+$ (lowering $S_-$), while $c_d^\dagger c_d = \frac{1}{2}$. In our case of no external magnetic field, $\epsilon_d = 0$, and Majorana fermion $\gamma_{d+} + \gamma_{d-} = (c_d + c_d^\dagger)/\sqrt{2}$ decouples from $H_x^{2\text{CK}}$ (while the other Majorana $\gamma_{d-} = i(c_d^\dagger - c_d)/\sqrt{2}$ participates in $H_x^{2\text{CK}}$). Namely, half of the impurity decouples from bath electrons, making 2CK a non-Fermi liquid. $H_x^{2\text{CK}}$ is diagonalized as $H_x^{2\text{CK}} = \sum_{c \geq 0} \epsilon_c^2 c_{2c}^\dagger c_{2c} + \sum_{k \geq k_p} \epsilon_k d_k^\dagger d_k + \text{(const.)}$, and its eigenstates are denoted as $|\tilde{E}_x^{2\text{CK}}\rangle$; $|\tilde{E}_x^{2\text{CK}}\rangle$ also has the other quantum numbers decoupled from $H_x^{2\text{CK}}$ such as total charge/spin and the charge difference.

We compute $S_{z,ii'} = \langle E_x^{2\text{CK}}|S_z|E_x^{2\text{CK}}\rangle$. The eigenstates $|\tilde{E}_x^{2\text{CK}}\rangle$ of 2CK connects with the eigenstates $|\tilde{E}_x^{2\text{CK}}\rangle = U_{\text{EK}}|\tilde{E}_x^{2\text{CK}}\rangle$ of the resonant-level model via Emery-Kivelson transformation $U_{\text{EK}}$ along Emery-Kivelson line [34]. Since $U_{\text{EK}} |\tilde{E}_x^{2\text{CK}}\rangle = S_z = c_d^\dagger c_d - 1/2 = \gamma_d + \gamma_{d-}$, $S_{z,ii'}$ is written as $S_{z,ii'} = i\langle \tilde{E}_x^{2\text{CK}}|\gamma_d + \gamma_{d-}|\tilde{E}_x^{2\text{CK}}\rangle$. After some calculations, we find

$$S_{z,ii'} = \frac{1}{2} \sum_{c', c'' \geq 0} B_{c', c''} \langle \tilde{E}_x^{2\text{CK}}|c_{c'} + c_{c''}^\dagger(c_{c''}^\dagger - c_{c''})|\tilde{E}_x^{2\text{CK}}\rangle$$

$$= \frac{1}{2} \sum_{c' \geq 0} B_{c', 0} \langle \tilde{E}_x^{2\text{CK}}|c_{c'}(c_{c'}^\dagger - c_{c'})|\tilde{E}_x^{2\text{CK}}\rangle.$$
where coefficient $B_{d±}$ connects $γ_{d±}$ and the excitation of $(c_{d−}^± ± c_{d+}−)/\sqrt{2}$ and $|(ℏ^2 E_{2CK}^±)(c_{d−}^± ± c_{d+}−)(c_{dν}^± − c_{dν}−)|² = 1$ or 0; for the detail of $B_{d±}$, see Ref. [34]. In the last equality, we used $B_{d±} = δ_{d0}$, coming from the decoupling of Majorana fermion $γ_{d+}$ from the bath. Since $B_{d+} − √T/T_{2CK}$ at $T ≪ T_{2CK}$, $S_{±γd±} ∝ √T/T_{2CK}$ in agreement with Eq. [S12].

We also compute $S_{−ii'} = (E_{1CK}^i)|S| E_{1CK}^i$. Using $U_{E}$ and $c_{d} = F_{d} S_{−i}$, where $F_{d}$ is a Klein factor, we have $S_{−ii'} = (E_{2CK}^i e^{−iφ/(0))F_{d} c_{d} E_{2CK}^i}$, where the boson field $φ/(0))$ results from the commutation between $S_{−}$ and $U_{E}$; see Ref. [34]. $φ/(0))$ and $F_{d}$ correspond to total charge degree of freedom. Here, $F_{d}$ is 1 or 0, hence, not related with $T/T_{2CK}$. And, $c_{d} = (c_{d+}− c_{d−})/2 + O(1/T_{2CK})$ does not provide $1/T/T_{2CK}$ in the leading order term. In contrast, $e^{−iφ/(0))}$ interestingly provides $−iφ/(0)) ∝ √T/T_{2CK}$, since the bosonic reservoir, included in the resonant-level model as being decoupled from $H_{2CK}^i$, also has the finite length of $~ hν_{F}/Δ ∼ hν_{F}/k_{B}T$. We show this, expanding $φ/(0))$ in terms of boson operators $b_{q}$ with momentum $q = n_q Δ/hν_{F}$. $φ/(0)) = ∑_{q>0} 1/n_q (b_{q}^† b_{q}) e^{−a_{q}/2}$, where $n_q$ is a positive integer and $a ∝ 1$ is the cutoff. Some calculations lead to

$$e^{−iφ/(0))} = ∑_{q>0} \left[ \exp \left( i \frac{1}{n_q} b_{q} e^{−a_{q}/2} \right) \exp \left( i \frac{1}{n_q} b_{q} e^{−a_{q}/2} \right) \exp \left( −i \frac{1}{2n_q} e^{−a_{q}} \right) \right].$$

The first and second terms in the squared bracket are $O(1)$ since $E_{1CK}^i$ are eigenstates of $b_{q}^† b_{q}$ with eigenvalues 0 or 1. Meanwhile, $π_{q>0} \exp \left( −i \frac{1}{2n_q} e^{−a_{q}} \right) = √1 − e^{−a_{q}/2} ∝ \sqrt{a_{q}} ∝ √T/T_{2CK}$ at $T ≪ T_{2CK}$. Hence, $S_{−ii'} ∝ √T/T_{2CK}$ is proved.

Next, we derive Eq. [S12] for 1CK. According to the bosonization and refermionization at Toulouse point [34], the spin degree of freedom of 1CK is also described by a similar resonant-level model, $H_{s1CK} = c_{d} : c_{d}^† c_{d} : + ∑_{k} c_{k} : c_{k}^† c_{k} : + \sqrt{Δ} ∑_{k} (c_{k}^† c_{k} + c_{k}^† c_{k}) = ∑_{α} c_{α}^† c_{α} : (const.),$ but with $Γ = k_{B}T_{1CK}$. Contrary to 2CK, it shows a Fermi liquid, and no Majorana fermion of the impurity decouples from the bath. We compute $S_{zii'} = (E_{1CK}^i |S| E_{1CK}^i)$, where $E_{1CK}^i$ is denoted the eigenstates of $H_{s1CK}$. Using another Emery-Kivelson transformation $U_{E}$, $E_{1CK}^i = U_{E}^† E_{1CK}^i U_{E} = U_{E}^† E_{1CK}^i$, we find $S_{zii'} = (E_{1CK}^i |S| E_{1CK}^i)$, $S_{zii'} = (E_{1CK}^i |S| E_{1CK}^i) − b_{ii'}′/2$, since $U_{E}^† S_{zii'} U_{E} = S_{z} = a_{d} = 1/2$. It is written as $S_{zii'} = ∑_{α} c_{α}^† c_{α} : + ∑_{α} c_{α}^† c_{α} :$. Since $B_{d} ∝ √T/T_{1CK}$ at $T ≪ T_{1CK}$ [34] and $E_{1CK}^i c_{d}^† c_{d}^† = 1$ or 0, we find $S_{zii'}' ∝ T/T_{1CK}$, in agreement with Eq. [S12]. Similarly, it is straightforward to show $S_{zii'} ∝ T/T_{1CK}$.

We also compute $S_{−ii'} = (E_{1CK}^i |S| E_{1CK}^i)$. The bosonization results in an expression similar to the 2CK, $S_{−ii'} = (E_{1CK}^i |S| E_{1CK}^i)$. It is however hard to handle $e^{i(\sqrt{2}−1)φ/(0))}$ with the irrational number $\sqrt{2}−1$ of Toulouse point. Instead, we study $S_{−ii'}$ using an effective theory near the strong-coupling fixed point [21]. At the fixed point, the Kondo singlet state decouples from Fermi-liquid excitations. Near the fixed point at $T ≪ T_{1CK}$, the singlet and the excitations are coupled, with coupling energy $∼ Δ/(N−1)/4 ∝ √T$. This modifies $E_{1CK}^i$ from $|E_{1CK}^{i0})$ as $|E_{1CK}^{i}) = |E_{1CK}^{i0}) + |δ_{i})$, where $|E_{1CK}^{i0})$'s are the states at the fixed point. The coupling energy leads to $|δ_{i}) ∝ T$, resulting in $S_{−ii'} ∝ T$, in agreement with Eq. [S12]. The same argument reproduces $S_{zii'} = (E_{1CK}^i |S| E_{1CK}^i) ∝ T/T_{1CK}$, which was obtained using the bosonization in the above.

### H. L dependence of $E_F$ for 1CK

Our numerical result of the $L$ dependence of $E_F$ at $T = 0$ and $L ≫ ξ_{1CK}$ in Eq. (4) is reproduced with the variational 1CK ground state by Yoshida [36],

$$|ψ_L⟩ = |ψ⟩|ψ⟩|ψ⟩|ψ⟩|ψ⟩/\sqrt{2}, \quad |ψ⟩|ψ⟩ = ψ^†|ψ⟩|ψ⟩, \quad ψ^†|ψ⟩|ψ⟩ = \frac{1}{\sqrt{N}} ∑_{k>0} e^{i k} |k⟩ + E_Y,$$

where $|ψ⟩ = \prod_{k≤k_F} e^{i k} |0⟩$ is the Fermi sea of the bath, $|0⟩$ is the vacuum state, $N$ is the normalization factor ensuring $⟨φ|ψ⟩|φ⟩|^2 = 1$, and $E_Y = D e^{−4/3hν_F}$ corresponds to $k_{B}T_{1CK}$; we here use $c_{k}^† = e_{α=1,k}^†$. This illustrates the Kondo singlet of the impurity spin and the electron spin created by $φ^†|ψ⟩|ψ⟩$. The spatial dependence of $φ^†|ψ⟩|ψ⟩$ is $φ(x) = \frac{1}{\sqrt{N}} ∑_{k>0} \sqrt{2} sin k x / ℏ ε_k + E_Y$, where $l → ∞$ is the total length of the one-dimensional bath and $E_Y ≡ hν_{F} k_{F} ≥ hν_{F}/ξ_{1CK}$. 


To study the $L$ dependence of $\mathcal{E}_F$, we compute $\text{Tr}_{x>L} \langle |\psi_Y\rangle \langle \psi_Y|$, by tracing out the states outside $L$. For this purpose, we decompose each single-electron operator,

$$c^\dagger_{k\sigma} = \sqrt{\frac{L}{T}} c^\dagger_{k\sigma,\text{in}} + \sqrt{1-\frac{L}{T}} c^\dagger_{k\sigma,\text{out}}, \quad \phi^\dagger_{Y\sigma} = \sqrt{1-p} \phi^\dagger_{Y\sigma,\text{in}} + \sqrt{p} \phi^\dagger_{Y\sigma,\text{out}},$$

where $c^\dagger_{k\sigma,\text{in(out)}} \sim \int_{x \leq L (x > L)} dx c^\dagger_{x\sigma} \sin kx$ creates an electron inside (outside) $L$ and $\phi^\dagger_{Y\sigma,\text{in(out)}} \sim \int_{x \leq L (x > L)} dx c^\dagger_{x\sigma} \phi_Y(x)$ ($c^\dagger_{x\sigma}$ creates a spin-$\sigma$ electron at $x$). $p = \int_L^\infty dx |\phi_Y(x)|^2 \approx 1/\pi k_L L \approx \xi_{1CK}/\pi L$ is the probability of finding the electron of $\phi_Y$ outside $L$. Accordingly, the Fermi sea is written as $|\mathcal{F}_{\text{S}}\rangle = \prod_{k \leq k_F} \left( \sqrt{\frac{L}{T}} c^\dagger_{k\sigma,\text{in}} + \sqrt{1-\frac{L}{T}} c^\dagger_{k\sigma,\text{out}} \right) |0\rangle_{\text{in(out)}} \approx \lim_{l \to \infty} |0\rangle_{\text{in(out)}}$, where $|0\rangle_{\text{in(out)}}$ denotes the vacuum state of $x \leq L$ ($x > L$) and $|\mathcal{F}_{\text{S}}\rangle_{\text{out}} = \prod_{k \leq k_F} c^\dagger_{k\sigma,\text{out}} |0\rangle_{\text{out}}$ is the Fermi sea outside $L$. Here, we used $l \gg L$, where the portion of plane waves inside $L$ can be ignored and $|\mathcal{F}_{\text{S}}\rangle_{\text{out}}$ is well defined. Using the decomposition, we find

$$|\psi_Y\rangle \approx \frac{1}{\sqrt{2}} \left( |\uparrow\rangle (\sqrt{1-p} \phi^\dagger_{\uparrow,\text{in}} + \sqrt{p} \phi^\dagger_{\uparrow,\text{out}}) - |\downarrow\rangle (\sqrt{1-p} \phi^\dagger_{\downarrow,\text{in}} + \sqrt{p} \phi^\dagger_{\downarrow,\text{out}}) \right) |0\rangle_{\text{in(out)}}.$$

Then, we compute $\text{Tr}_{x>L} \langle \psi_Y | \psi_Y \rangle = \sum_i \langle \psi_{i,\text{out}} | \psi_Y \rangle \langle \psi_Y | \psi_{i,\text{out}} \rangle$, where $\psi_{i,\text{out}}$'s are relevant states outside $L$, $|\psi_{i,\text{out}}\rangle \in \{|\mathcal{F}_{\text{S}}\rangle_{\text{out}}, \phi^\dagger_{\uparrow,\text{out}} |\mathcal{F}_{\text{S}}\rangle_{\text{out}}, \phi^\dagger_{\downarrow,\text{out}} |\mathcal{F}_{\text{S}}\rangle_{\text{out}} \}$. The result is

$$\text{Tr}_{x>L} \langle \psi_Y | \psi_Y \rangle \simeq (1-p) \langle \psi_{1,\text{in}} | \psi_{1,\text{in}} \rangle + \frac{p}{2} \langle \psi_{2,\text{in}} | \psi_{2,\text{in}} \rangle + \frac{p}{2} \langle \psi_{3,\text{in}} | \psi_{3,\text{in}} \rangle,$$

where $|\psi_{1,\text{in}}\rangle = \frac{1}{\sqrt{2}} (|\uparrow\rangle \phi^\dagger_{\uparrow,\text{in}} - |\downarrow\rangle \phi^\dagger_{\downarrow,\text{in}}) |0\rangle_{\text{in}}$, $|\psi_{2,\text{in}}\rangle = |\downarrow\rangle |0\rangle_{\text{in}}$, and $|\psi_{3,\text{in}}\rangle = |\uparrow\rangle |0\rangle_{\text{in}}$.

We calculate $\mathcal{E}_F(\text{Tr}_{x>L} \langle \psi_Y | \psi_Y \rangle)$, using a witness operator similar to Eq. (6),

$$X_Y = \frac{2}{\log 2} |\psi_{1,\text{in}}\rangle \langle \psi_{1,\text{in}}| - \left( \frac{2}{\log 2} - 1 \right) I_{1,\text{in}},$$

where $I_{1,\text{in}} = \sum_{\eta=\uparrow,\downarrow,\sigma=\uparrow,\downarrow} |\eta\rangle \langle \eta| \otimes \phi^\dagger_{Y\sigma,\text{in}} \langle \phi_{Y\sigma,\text{in}}|$ and $\phi^\dagger_{Y\sigma,\text{in}} = \phi^\dagger_{Y\sigma,\text{in}} |0\rangle_{\text{in}}$. This operator is the optimal witness operator for $\mathcal{E}_F(|\psi_{i,\text{in}}\rangle)$ with $i = 1, 2, 3$, namely, it provides the exact value of $\mathcal{E}_F(|\psi_{i=1,2,3,\text{in}}\rangle)$; one checks $\mathcal{E}_F(|\psi_{1,\text{in}}\rangle) = \langle \psi_{1,\text{in}} | X_Y | \psi_{1,\text{in}} \rangle = 1$, $\mathcal{E}_F(|\psi_{2,\text{in}}\rangle) = \mathcal{E}_F(|\psi_{3,\text{in}}\rangle) = 0$. According to the duality [11][12] between Eqs. (1) and (2), the expectation value of $X_Y$ equals the exact value of $\mathcal{E}_F$ for any mixture of $|\psi_{i=1,2,3,\text{in}}\rangle$ including $\mathcal{E}_F(\text{Tr}_{x>L} \langle \psi_Y | \psi_Y \rangle)$. We obtain $\mathcal{E}_F(\text{Tr}_{x>L} \langle \psi_Y | \psi_Y \rangle) = \text{Tr} X_Y \text{Tr}_{x>L} \langle \psi_Y | \psi_Y \rangle \approx 1 - p$, namely, $1 - \mathcal{E}_F(\text{Tr}_{x>L} \langle \psi_Y | \psi_Y \rangle) \approx p \propto \xi_{1CK}/L$. This confirms the universal power law in Eq. (4), which we numerically find in the main text. This computation based on $X_Y$ indicates the usefulness of witness operators for analytically studying macroscopic entanglement $\text{EoF}$ in many-body mixed-states.

I. Kondo cloud at finite temperature

In Fig. [S4] we present our numerical result of the dependence of $\mathcal{E}_F(\rho)$ on $L$ at finite $T$. Figure [S4] shows that as $L$ decreases, $\mathcal{E}_F$ starts to decrease near $L \approx 1.2 \xi_{1CK}$ at $T \lesssim T_{1CK}$, while roughly near thermal length $L \approx L_T \approx \hbar v_F/k_B T$ at $T \gtrsim T_{1CK}$. This means that the size of Kondo cloud is $\xi_{1CK}$ and robust against thermal effects at $T \lesssim T_{1CK}$, while it is roughly $L_T$, decreasing with increasing $T$, at $T \gtrsim T_{1CK}$. Moreover, Fig. [S5] suggests that the two 1CK power-law decays in Eqs. (3) and (4) are additive at $T \ll T_{1CK}$ and $L \gg 1CK$,

$$\mathcal{E}_F \simeq 1 - a_1 \left( \frac{T}{T_{1CK}} \right)^2 - b_1 \left( \frac{\xi_{1CK}}{L} \right).$$

Together with the fact that the two 1CK power laws are not connected by the usual replacement of $k_B T \leftrightarrow \hbar v_F / L$ by the uncertainty relation (as their power-law exponents are different), these unusual findings indicate that the mechanism of entanglement suppression by thermal effects differs from that by the partial trace over $x > L$. Note that we are unable to definitely conclude whether the cloud size is $L_T$ at $T \gtrsim T_{1CK}$, because the numerical results of the upper and lower bounds of $\mathcal{E}_F$ are not close enough to each other; the witness operator $X$ is devised from the entanglement feature of the ground and low-energy eigenstates, hence, less efficient at $T \gtrsim T_{1CK}$ or $L \lesssim \xi_{1CK}$.

All these findings can be understood by the following argument. At finite $T$, $\mathcal{E}_F(\rho)$ is mainly contributed by the excited states $|E_i\rangle = b_{i\uparrow} |\uparrow\rangle |\epsilon_{i\uparrow}\rangle + b_{i\downarrow} |\downarrow\rangle |\epsilon_{i\downarrow}\rangle$ of $E_i \sim k_B T$. They have $\mathcal{E}_F(|E_i\rangle) \simeq 1 - 2 |S_{z,i}|^2 / \log 2 \approx 1 - (b_{i\uparrow})^2 -$
FIG. S4. Kondo cloud at finite temperature. Dependence of $\mathcal{E}_F$ on $L$ at different $T$'s, $T/T_{1CK} = 0, 10^{-3/4}/4, 1, 10$; the results of $T/T_{1CK} = 0$ and $10^{-3/4}/4$ are almost overlapped. This shows that the cloud size is about $\xi_{1CK}$ at $T \lesssim T_{1CK}$, while it decreases at $T \gtrsim T_{1CK}$ as $T$ increases. Empty (filled) symbols represent a upper (lower) bound of $\mathcal{E}_F$.

FIG. S5. Dependence of $\mathcal{E}_F(T = 0) - \mathcal{E}_F(T)$ on $L$ at different $T$'s, $T/T_{1CK} = 10^{-3/4}/4, 10^{-5/4}/4, 10^{-7/4}/4$. This shows that $\mathcal{E}_F(T = 0) - \mathcal{E}_F(T)$ is almost independent of $L$ at $T \ll T_{1CK}$ and $L \gg \xi_{1CK}$. Empty (filled) symbols represent a upper (lower) bound of $\mathcal{E}_F$.

$|b_{ij}|^2/(2 \log 2)$. For larger $E_i$, $|S_{z,ii}|^2 \propto (|b_{ij}|^2 - |b_{ij}|^2)^2$ increases, as $|E_i|$ more deviates from the exact Bell state. Our numerical results imply that the dependence of $\mathcal{E}_F$ on $T$ reflects this behavior, hence, the entanglement of excited states $|E_i\rangle$. On the other hand, the $L$ dependence of $\mathcal{E}_F$ is related to the loss of the wave functions of $|e_{\uparrow R}\rangle$ and $|e_{\downarrow L}\rangle$ by the partial trace over $x > L$. At $T \ll T_{1CK}$ and $L \gg \xi_{1CK}$, the two mechanisms ($|S_{z,ii}|^2$ and the partial wave-function loss) seem to work independently, resulting in the additive scaling law in Eq. (S17) as $\mathcal{E}_F \simeq (1 - a_1(T/T_{1CK})^2)(1 - b_1\xi_{1CK}/L) \simeq 1 - a_1(T/T_{1CK})^2 - b_1\xi_{1CK}/L$. The size of Kondo cloud, measured by $\mathcal{E}_F$, may directly reflect the spatial extension of the wave functions $\langle x|e_{\uparrow R}\rangle$ and $\langle x|e_{\downarrow L}\rangle$ participating in excited-state entanglement.