This is the accepted manuscript made available via CHORUS. The article has been published as:

**Uniform Additivity in Classical and Quantum Information**
Andrew Cross, Ke Li, and Graeme Smith
DOI: [10.1103/PhysRevLett.118.040501](https://doi.org/10.1103/PhysRevLett.118.040501)
Uniform Additivity in Classical and Quantum Information

Andrew Cross,1 Ke Li,1,2,3 and Graeme Smith4

1IBM TJ Watson Research Center, Yorktown Heights, NY 10598, USA
2Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
3IQIM, California Institute of Technology, Pasadena, CA 91125, USA
4JILA and Department of Physics, University of Colorado, Boulder, CO 80309, USA

(Dated: November 21, 2016)

Information theory quantifies the optimal rates of resource interconversions, usually in terms of entropies. However, nonadditivity often makes evaluating entropic formulas intractable. In a few auspicious cases, additivity allows a full characterization of optimal rates. We study uniform additivity, which is easily evaluated and captures all known additive quantum formulas. Our complete characterization of uniform additivity exposes an intriguing new additive quantity and identifies a remarkable coincidence—the classical and quantum uniformly additive functions with one auxiliary variable are identical.

Entropies tell us how much information is stored in a system. As a result, the answers to information theoretic questions are usually found in terms of entropies evaluated on systems arising in optimal protocols. For example, the communication capacity of a classical channel \( \mathcal{N} \) that maps random variable \( X \) to \( Y \) is given by the maximization \( C(\mathcal{N}) = \max_X I(X; Y) \), where the mutual information \( I(X; Y) = H(X) + H(Y) - H(XY) \) is a linear combination of entropies [25]. Similarly, the cost of transmitting a quantum state \( \rho \) on system \( A \) is its von Neumann entropy \( H(A) = -\text{tr} \rho_A \log \rho_A \). A noisy quantum communication channel \( \mathcal{N} : A \rightarrow B \) can be mathematically extended to an isometry \( U : A \rightarrow BE \) of the input with an independent and inaccessible environment. Such a channel can be applied to a state \( \phi_{V_1...V_n} \) to create a state \( \rho_{VBE} \). More generally, \( V \) may have many subsystems, and we may use \( \phi_{V_1...V_n} \) to create \( \rho_{V_1...V_nBE} \). We can use such a state to generate an entropic formula: \( f_\alpha(U_N) = \max_{\phi_{V_1...V_n}} f_\alpha(U_N; \phi_{V_1...V_n}) \) with \( f_\alpha(U_N; \phi_{V_1...V_n}) = \sum_{s \in \mathcal{P}(V_1...V_nBE)} \alpha_s H(\rho_s) \), where \( \mathcal{P}(V_1...V_nBE) \) ranges over all collections of subsystems from \( V_1...V_nBE \), and \( H(\rho_s) \) is the entropy of collection \( s \). We call the \( V_1...V_n \) systems auxiliary variables, and they can a priori have arbitrary, even infinite, dimensions. Most operationally relevant quantities in quantum information can be expressed as a regularization of such a formula:

\[
 f_\alpha^{\infty}(\mathcal{N}) = \lim_{n \rightarrow \infty} \frac{1}{n} f_\alpha \left( \mathcal{N}^{\otimes n} \right),
\]

where \( \mathcal{N}^{\otimes n} \) is the \( n \)-fold parallel use of channel \( \mathcal{N} \). The auxiliary variables in an entropic formula are usually related operationally to the structure of optimal protocols; for example, the optimal distribution \( X \) that maximizes \( C(\mathcal{N}) = \max_X I(X; Y) \) to give the classical capacity defines a distribution of capacity-achieving error correcting codes.

The infinite-dimensional optimization of Eq.(1), which is called a multi-letter formula, is usually intractable. In some rare cases additivity allows a substantial simplification. An entropic formula \( f_\alpha(U_N) \) is additive if \( f_\alpha(U_N \otimes U_M) = f_\alpha(U_N) + f_\alpha(U_M) \) for all channels \( \mathcal{N} \) and \( \mathcal{M} \). When \( f_\alpha \) is additive, we have \( f_\alpha^{\infty}(U_N) = f_\alpha(U_N) \), which is called a single-letter formula. There are single-letter formulas for the classical capacity of a classical channel [1], the entanglement-assisted capacity of a quantum channel [2], and the quantum capacity of a quantum channel with access to a special zero-capacity assistance channel[3]. Furthermore, there are single-letter formulas for the classical capacity of an entanglement breaking channel [4] and the quantum capacity of degradable channels [5]. A single-letter formula often leads to a tractable means of evaluating a quantity, allowing us to completely characterize the optimal performance for information transmission and storage.

Many relevant entropic formulas are nonadditive, especially in the quantum setting[6–10]. Optimal performance is thus captured only by a multi-letter formula, which is intractable to evaluate. Even the capacities themselves can exhibit nonadditivity, displaying fundamentally quantum synergies not present classically [9–13]. As a result, many basic questions in quantum information theory remain open—the classical and quantum capacities of most channels are unknown, and even deciding if a quantum channel has nonzero quantum capacity...
there must be a pair of states \( \tilde{\phi} \) that clarifies when and where to expect quantum synergies among new classes of channels with additive capacities, and the method opens a line of attack on a variety of questions, from classical multiuser information theory to finding new classes of channels with additive capacities, and clarifies when and where to expect quantum synergies like superactivation [11].

Entropy inequalities express relationships between entropies of different collections of subsystems that are satisfied for all states. Subadditivity of entropy, for example, tells us that \( H(A) + H(B) - H(AB) \geq 0 \), or equivalently \( I(A;B) \geq 0 \). Its generalization, strong subadditivity [15], tells us that conditional mutual information is also positive:

\[
I(A;B|C) = H(AC) + H(BC) - H(ABC) - H(C) \geq 0.
\]

The set of \( (2^n - 1) \)-dimensional entropy vectors \( \mathbf{v} = (H(X_1),\ldots,H(X_n)) \) that can be realized by classical probability distributions on \( X_1\ldots X_n \) form a cone, whose study in terms of linear programming was formalized in [16]. The larger cone of realizable quantum entropies was studied in [17]. Entropy inequalities are the key to proving additivity when it exists.

If \( f_\alpha \) is an additive formula with one auxiliary variable [26], for any pair of channels \( \mathcal{N}, \mathcal{M} \) and any state \( \phi_{V,A_1,A_2} \), there must be a pair of states \( \tilde{\phi}_{V,A_1} \) and \( \tilde{\phi}_{V,A_2} \) such that

\[
f_\alpha(\mathcal{N} \otimes \mathcal{M}, \phi_{V,A_1,A_2}) \leq f_\alpha(\mathcal{N}, \tilde{\phi}_{V,A_1}) + f_\alpha(\mathcal{M}, \tilde{\phi}_{V,A_2}).
\]

We call such a mapping \( \phi_{V,A_1,A_2} \to (\tilde{\phi}_{V,A_1},\tilde{\phi}_{V,A_2}) \) a decoupling. In principle, the appropriate decoupling may depend in an arbitrary way on the channels \( \mathcal{N}, \mathcal{M} \) and the state \( \phi_{V,A_1,A_2} \). In practice, useful decouplings are invariably what we call standard decouplings, which have a very simple form and are described in Fig. 2. Once we have fixed a decoupling and \( f_\alpha \), we can use entropy inequalities to determine if Eq. (3) is satisfied. When \( f_\alpha \) does satisfy Eq. (3) with \( (\tilde{\phi},\tilde{\phi}) \) defined by a standard decoupling \( D \), we say \( f_\alpha \) is uniformly subadditive with respect to \( D \). Since we also have \( f_\alpha(\mathcal{N} \otimes \mathcal{M}, \tilde{\phi} \otimes \tilde{\phi}) = f_\alpha(\mathcal{N}, \tilde{\phi}) + f_\alpha(\mathcal{M}, \tilde{\phi}) \), subadditivity implies that

\[
f_\alpha(\mathcal{N} \otimes \mathcal{M}) = f_\alpha(\mathcal{N}) + f_\alpha(\mathcal{M}) \tag{4}
\]

and we call \( f_\alpha \) uniformly additive with respect to \( D \). All known proofs of quantum additivity proceed by choosing a standard decoupling and proving Eq. (3) via entropy inequalities [2, 3, 18].

We have found all entropic formulas \( f_\alpha \) that are uniformly additive with respect to standard decouplings. We do this by enumerating all standard decouplings, and using the linear programming formulation of entropy inequalities to determine which \( f_\alpha \) are uniformly subadditive for each decoupling. Our approach captures all previously known examples of additive formulas and more. This method opens a line of attack on a variety of questions, from classical multiuser information theory to finding new classes of channels with additive capacities, and clarifies when and where to expect quantum synergies like superactivation [11].

Formulas with no auxiliary variables are particularly simple:

\[
f_\alpha(U^n_A, \phi_A) = \alpha_B H(B) + \alpha_E H(E) + \alpha_{BE} H(BE). \tag{5}
\]

Here we have only one standard decoupling to consider: \( \phi_{A_1,A_2} \to (\phi_{A_1},\phi_{A_2}) \). The conditions for uniform additivity in this case are...
\[ a_B + a_{BE} \geq 0, \quad a_E + a_{BE} \geq 0, \quad (6) \]
\[ a_B + a_E + a_{BE} \geq 0, \quad a_{BE} \geq 0. \]

These inequalities define a cone of \( a \)s, which we refer to as a uniform additivity cone. Eq. (6) describes this cone in terms of its facets, but a cone can equally well be described in terms of extremal rays: letting

\[ \alpha_0 = (1, 0, 0) \equiv H(B), \quad \alpha_1 = (0, 1, 0) \equiv H(E) \quad \text{(7)} \]
\[ \alpha_2 = (0, -1, 1) \equiv H(B|E), \quad \alpha_3 = (-1, 0, 1) \equiv H(E|B), \]

so that \( f_\alpha \) is uniformly additive with respect to the standard decoupling exactly when \( \forall N, M, \phi_{A_1 A_2} \) we have \( \Delta(\alpha, U_N \otimes U_M, \phi_{A_1 A_2}) \geq 0 \). We make use of the alternate characterization of Eq. (6) in terms of extremal rays, Eq. (7). It is easy to verify that the \( a \)s associated with each of the extremal rays \( H(B), H(E), H(E|B), \) and \( H(B|E) \) lead to positive \( \Delta(\alpha, U_N \otimes U_M, \phi_{A_1 A_2}) \). For example, \( H(B) \) corresponds to \( (\alpha_B, \alpha_E, \alpha_{BE}) = (1, 0, 0) \) and \( \Delta(\alpha, U_N \otimes U_M, \rho_{A_1 A_2}) = I(B_1; B_2) \geq 0 \), while \( H(B|E) \) corresponds to \( (\alpha_B, \alpha_E, \alpha_{BE}) = (0, -1, 1) \) and gives

\[ \Delta(\alpha, U_N \otimes U_M, \rho_{A_1 A_2}) = I(B_1 E_1; B_2 E_2) - I(E_1; E_2), \]

which is also positive for all \( \rho_{A_1 A_2}, H(E) \) and \( H(E|B) \) follow mutatis mutandis. Eq. (6) is thus a sufficient condition for uniform additivity. To see that it is also a necessary condition, we find states (in fact, classical distributions) \( p^0, p^1, p^2, p^3 \) and channels \( N, M \) such that

\[ \Delta(\alpha, U_N \otimes U_M, p^0) = \alpha_B + \alpha_{BE} \]
\[ \Delta(\alpha, U_N \otimes U_M, p^1) = \alpha_E + \alpha_{BE} \]
\[ \Delta(\alpha, U_N \otimes U_M, p^2) = \alpha_B + \alpha_E + \alpha_{BE} \]
\[ \Delta(\alpha, U_N \otimes U_M, p^3) = \alpha_{BE}. \]

This shows that for any \( \alpha \) that doesn’t satisfy Eq. (6) there are states and channels such that \( \Delta(\alpha, U_N \otimes U_M, p) < 0 \). Thus, Eq. (6) are both necessary and sufficient for uniform additivity.

Formulas with one auxiliary variable require us to consider multiple decouplings, capturing the choice of \( \tilde{V} \) and \( \tilde{V} \) in the decoupling map \( D : \psi_{V A_1 A_2} \rightarrow (\tilde{\psi}_{\tilde{V} A_1}, \tilde{\psi}_{\tilde{V} A_2}) \). A standard decoupling always has \( \tilde{V} = M_2 V \) with \( M_2 \) chosen from \( \{ \emptyset, B_2, E_2, B_2 E_2 \} \) and \( \tilde{V} = M_1 V \) with \( M_1 \) chosen from \( \{ \emptyset, B_1, E_1, B_1 E_1 \} \). We can parametrize these by \( (a, b) \), with \( a \) and \( b \) running from 0 to 3. We take advantage of two simplifications that can be made without loss of generality. First, given \( f_\alpha, \alpha = (\alpha^a, \alpha^b) \) with \( \alpha^a = (\alpha_B, \alpha_E, \alpha_{BE}) \) and \( \alpha^b = (\alpha_V, \alpha_{BV}, \alpha_{EV}, \alpha_{BEV}) \), we can define \( f^a_\alpha \) and \( f^b_\alpha \) such that \( f_\alpha \) is uniformly additive with respect to decoupling \( (a, b) \) if and only if \( f^a_\alpha \) is uniformly additive with respect to the decoupling \( \phi_{A_1 A_2} \rightarrow (\tilde{\phi}_{\tilde{V} A_1}, \tilde{\phi}_{\tilde{V} A_2}) \) and \( f^b_\alpha \) is uniformly additive with respect to \( (a, b) \). Second, these formulas have two useful symmetries that reduce the number of decouplings we must consider: 1) for any additive formula, we get a similar additive formula by exchanging \( B \) and \( E \) and 2) \( f^b_\alpha \) with \( \alpha^b = (\alpha_V, \alpha_{BV}, \alpha_{EV}, \alpha_{BEV}) \) is equivalent via purification of the quantum state to \( f^b_\alpha \) with \( \tilde{\alpha}^b = (\alpha_{BEV}, \alpha_{EV}, \alpha_{BY}, \alpha_V) \). This leaves only 5 inequivalent decouplings to be considered.

Table I describes the functions \( f^b_\alpha \) that are uniformly additive with respect to the 5 inequivalent decouplings. They are positive linear combinations\([27]\) of the extreme rays in the corresponding row of the table. The uniformly additive functions with respect to decoupling \( (a, b) \) are the sum of any \( f^a_\alpha \) satisfying Eq. (6) and such an \( f^b_\alpha \) found from Table I.

We find many familiar additive quantities in this way. For example, maximum output entropy \((\max_{\phi_{A}} H(B))\) satisfies Eq. (6). The quantity \(-H(B|V)\) was shown to be additive in [18], and later referred to as reverse coherent information\([19]\). Since \( H(B) \) satisfies Eq. (6) and \(-H(B|V)\) is uniformly additive with respect to multiple decouplings, so is their sum \( H(B) - H(B|V) = I(B; V) \), whose maximization gives the entanglement assisted assisted capacity.

One extreme ray of the \((1, 2)\) decoupling’s additive
<table>
<thead>
<tr>
<th>case</th>
<th>$(a,b)$</th>
<th>$M_1$</th>
<th>$M_2$</th>
<th>equivalents</th>
<th>Additive Cone</th>
<th>Extreme Rays</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>(3,3)</td>
<td>$B_1E_1$</td>
<td>$B_2E_2$</td>
<td>(0,0)</td>
<td>$\alpha_V + \alpha_{BV} + \alpha_{EV} \geq 0, \alpha_V + \alpha_{EV} \geq 0, \alpha_{BV} \geq 0$</td>
<td>$-H(E</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$-H(B</td>
</tr>
<tr>
<td>2.</td>
<td>(3,2)</td>
<td>$B_1E_1$</td>
<td>$E_2$</td>
<td>(2,3), (3,1), (1,3), (1,0) (0,1), (2,0), (0,2)</td>
<td>$\alpha_{BV} \leq 0, \alpha_V + \alpha_{BV} \geq 0$</td>
<td>$-H(BE</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$-H(B</td>
</tr>
<tr>
<td>3.</td>
<td>(3,0)</td>
<td>$B_1E_1$</td>
<td>$\emptyset$</td>
<td>(0,3)</td>
<td>$\alpha_{EV} \leq 0, \alpha_{BV} \leq 0$</td>
<td>$H(E</td>
</tr>
<tr>
<td>4.</td>
<td>(1,1)</td>
<td>$B_1$</td>
<td>$B_2$</td>
<td>(2,2)</td>
<td>$\alpha_{EV} = 0, \alpha_V \geq 0, \alpha_{BEV} \geq 0$</td>
<td>$-H(B</td>
</tr>
<tr>
<td>5.</td>
<td>(1,2)</td>
<td>$B_1$</td>
<td>$E_2$</td>
<td>(2,1)</td>
<td>$\alpha_{BEV} \geq 0, \alpha_V \geq 0$</td>
<td>$\pm H(E</td>
</tr>
</tbody>
</table>

| Extreme Rays | $H(E|BV), -H(E|V)$ | $H(E|BV), -H(E|V), \pm H(BE|V)$ | $-H(B|V), H(E|BV)$ | $\pm H(E|V) - H(B|V)$ |

TABLE I: Functions $f^c_\alpha$, that are uniformly subadditive with respect to the 5 inequivalent standard decouplings. Fixing a decoupling $D$, a single auxiliary variable $f_\alpha$ is uniformly subadditive with respect to $D$ exactly when it can be written as a sum of $f_{\alpha}^{\emptyset}$ satisfying Eq.6 and $f^c_\alpha$, that is a positive linear combination of the extreme rays in the row corresponding to $D$. Multiple auxiliary variables are all found similarly.

The cone is particularly intriguing:

$$I^{cc}(N) = \max_{\phi_{VA}} [H(VB) - H(VE)].$$

We call this quantity the completely coherent information, since its relationship to the coherent information $I^{coh}(N) = \max_{A} [H(B) - H(E)]$ is similar to the relationship between completely positive and positive maps. The version of this quantity evaluated on states was shown in [20] to be a lower bound on the communication cost of exchanging the $B$ and $E$ systems, but it was not realized that it is additive. We also show that $I^{cc}$ is also an upper bound for the jointly achievable quantum communication rate from $A$ to either $B$ or $E$. We have not found a clear operational interpretation of this quantity.

We now consider formulas with multiple auxiliary variables. For concreteness, suppose we have some formula depending on two auxiliary variables $V_1$ and $V_2$. A standard decoupling is a mapping from a state $\phi_{V_1V_2A_1A_2}$ to two states $\phi_{V_1V_2A_1A_2}$ and $\phi_{V_1V_2A_1A_2}$ that we get by choosing to incorporate (or not) $B_2$ and $E_2$ into one of $V_1$ and $V_2$ (and similarly for $B_1$, $E_1$ in $V_1$ and $V_2$). Since $V_1$ and $V_2$ should be non-overlapping, it is necessary to impose some consistency on the decouplings $(a_1,b_1)$ and $(a_2,b_2)$. These also give rise to a third decoupling, which we call $(a^*,b^*)$, that tells us which systems get included in the joint systems $V_1V_2$ and $V_1V_2$.

In this case it is possible to separate the variables much as we did in the single-variable case. Indeed, any $f_\alpha$ with $\alpha = (\alpha^{\emptyset}, \alpha^{V_1}, \alpha^{V_2}, \alpha^{V_1V_2})$ [28] is uniformly additive with respect to decoupling $\{(a_1,b_1), (a_2,b_2)\}$ exactly when $f^{\emptyset}_\alpha$, $f^{V_1}_\alpha$, $f^{V_2}_\alpha$, and $f^{V_1V_2}_\alpha$ are uniformly additive with respect to their respective decouplings. The same is true for more auxiliary variables. For any number of auxiliary variables, all $f_\alpha$ uniformly additive with respect to standard decouplings can be constructed from Table I and Eq. (6) [21].

Surprisingly, carrying out the same analysis as above for classical states and channels yields exactly the same set of uniformly additive functions for one auxiliary variable. This is in spite of the fact that the classical and quantum entropy cones do not coincide. This coincidence of uniformly additive functions may explain a well-known phenomenon: Formulas that solve classical information theory problems often tend to have corresponding quantum formulas that solve an appropriately coherized version of the problem [29]. In these cases, the classical and quantum problems have a solution for the same reason: the existence of an appropriately additive formula whose additivity proofs are formally equivalent. It would be very nice to formalize this apparent correspondence and explore its limits.

In some cases, nonadditive formulas can become additive when evaluated on special classes of channels. For example, while both the holevo information and minimum output entropy are nonadditive [8], for entanglement breaking channels they become additive. Similarly, while coherent information is nonadditive [6], it is additive on degradable channels [5]. Understanding such conditional additivities is an important open questions, and we are currently exploring the application of our techniques to finding special classes of channels that have additive capacities. We have identified a new criterion for the additivity of coherent information: informational degradability. We say a channel is informationally degradable if for any input state $\phi_{VA}$ we have $I(V;B) \geq I(V;E)$. This class includes degradable channels. We suspect informational degradability is the only single-letter entropic constraint on a channel that implies this additivity. We have also found a set of entropic constraints that imply a state is of the c-q form, which should be useful for studying classical and private capacities of quantum channels.

We have identified the limits of the techniques used in all known instances of quantum additivity. There are some classical formulas that are additive but not uniformly additive (e.g., minimum output entropy of a classical channel). Proving additivity in these cases requires...
knowledge of the optimizing state (in the case of minimum output entropy of a quantum channel, the optimal state is a pure state, which for classical channels is also a product state.). One potential path to new quantum additive formulas beyond what we have found is to better understand the optimizing state in an entropic formula. At this point we know of no examples where this can be done, but they may well exist.

ACKNOWLEDGEMENTS

KL acknowledges NSF grants CCF-1110941 and CCF-111382 and GS acknowledges NSF grant CCF-1110941.

[21] See supplemental material, which includes Refs. [22-24].
[25] The entropy of a random variable $X$ is $H(X) = -\sum_x p_x \log p_x$.
[26] We focus on 1 auxiliary variable for simplicity. Multiple variables can be handled similarly.
[27] i.e., linear combinations with positive coefficients
[28] Here $a^V = (a_B, a_E, a_{BE})$,

$$a^V_1 = (a_{V_1}, a_{BV_1}, a_{EV_1}, a_{BEV_1}),$$

$$a^V_2 = (a_{V_2}, a_{BV_2}, a_{EV_2}, a_{BEV_2}),$$

and $a^{V_1V_2} = (a_{V_1V_2}, a_{BV_1V_2}, a_{EV_1V_2}, a_{BEV_1V_2})$.
[29] Examples of this include 1) the correspondence between classical capacity of a classical channel and the entanglement assisted capacity of a quantum channel, 2) the connection between Slepian-Wolf and state merging, and 3) the correspondence between Csiszar-Korner solution to the broadcast channel with confidential messages and the recent analysis of the quantum one-time pad.