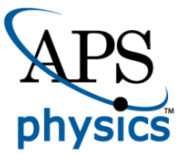


Quantum and Private Capacities of Low-Noise Channels

Felix Leditzky, Debbie Leung, and Graeme Smith

Phys. Rev. Lett. **120**, 160503 — Published 20 April 2018

DOI: 10.1103/PhysRevLett.120.160503

Quantum and private capacities of low-noise channels

Felix Leditzky,[1,2,*] Debbie Leung,[3,†] and Graeme Smith[1,2,4,‡]

[1]*JILA, University of Colorado/NIST, 440 UCB, Boulder, CO 80309, USA*
[2]*Center for Theory of Quantum Matter, University of Colorado, Boulder, Colorado 80309, USA*
[3]*Institute for Quantum Computing, University of Waterloo, Waterloo, Ontario, Canada. N2L 3G1.*
[4]*Department of Physics, University of Colorado, 390 UCB, Boulder, CO 80309, USA*

(Dated: March 22, 2018)

We determine both the quantum and the private capacities of low-noise quantum channels to leading orders in the channel's distance to the perfect channel. It has been an open problem for more than 20 years to determine the capacities of some of these low-noise channels such as the depolarizing channel. We also show that both capacities are equal to the single-letter coherent information of the channel, again to leading orders. We thus find that, in the low noise regime, super-additivity and degenerate codes have negligible benefit for the quantum capacity, and shielding does not improve the private capacity beyond the quantum capacity, in stark contrast to the situation when noisier channels are considered.

Any point-to-point communication link can be modeled as a quantum channel $\mathcal{N}$ from a sender to a receiver. Of fundamental interest are the *capacities* of $\mathcal{N}$ to transmit data of various types such as quantum, private, or classical data. Informally, the capacity of $\mathcal{N}$ to transmit a certain type of data is the optimal rate at which that data can be transmitted with high fidelity given an asymptotically large number of uses of $\mathcal{N}$. Capacities of a channel quantify its value as a communication resource.

In the classical setting, the capacity of a classical channel $\mathcal{N}$ to transmit classical data is given by Shannon's noisy coding theorem [1]. While operationally, the capacity-achieving error correcting codes may have increasingly large block lengths, the capacity can be expressed as a *single letter formula*: it is the maximum *mutual information* that a *single* channel use can generate between the input and output distributions.

The quantum capacity. In the quantum setting, to every quantum channel $\mathcal{N}$, one can associate an *environment* to which information is leaked when the channel $\mathcal{N}$ is used. The (1-shot) coherent information $I_c(\mathcal{N})$ of $\mathcal{N}$ is defined as the maximum of input-output mutual information less the input-environment mutual information. Intuitively, this quantifies a channel's ability to transmit quantum data to the receiver while minimizing the information leaked to the environment, as required by the well-known quantum principle that information gain in the environment implies disturbance of transmitted data. (See supplementary material [2] for details.) The Lloyd-Shor-Devetak (LSD) theorem [3–5] for the capacity of $\mathcal{N}$ to transmit quantum data, denoted $Q(\mathcal{N})$, makes this precise: they prove that the capacity is given by the expression $Q(\mathcal{N}) = \lim_{n\to\infty} \frac{1}{n} I_c(\mathcal{N}^{\otimes n})$. Here, $I_c(\mathcal{N}^{\otimes n})$ is the coherent information evaluated on n copies of $\mathcal{N}$, and is called the *n-shot coherent information* of $\mathcal{N}$. For special channels called *degradable* channels, the coherent information is *weakly additive*, meaning that $I_c(\mathcal{N}^{\otimes n}) = nI_c(\mathcal{N})$ [6], hence the capacity is the 1-shot coherent information and can be evaluated in principle. In general, the coherent information can be *superadditive*, $I_c(\mathcal{N}^{\otimes n}) > nI_c(\mathcal{N})$, and the potentially unbounded optimization over n in the capacity expression is necessary [7]. Twenty years after the LSD theorem was found, there is still no known algorithm to compute the capacity of a given channel. Furthermore, the n-shot coherent information can be positive for some small n while the 1-shot coherent information is zero [7]. Moreover, given any n, there is a channel whose n-shot coherent information is zero but whose quantum capacity is positive [8]. Thus we do not have a general method to determine if a given channel has positive quantum capacity.

Even for the qubit depolarizing channel, which is the quantum analogue of the binary symmetric channel and acts as $\mathcal{D}_p(\rho) = (1 - \frac{4p}{3})\rho + \frac{4p}{3}\frac{I}{2}$, our understanding of the quantum capacity is limited. The perfect channel ($p = 0$) has unit capacity, while a no-cloning argument gives $Q(\mathcal{D}_p) = 0$ for $p \geq 1/4$ [9]. However, $Q(\mathcal{D}_p)$ is unknown otherwise, despite substantial effort (see e.g. [10–12]). For $p \approx 0.2$, the 1-shot coherent information vanishes, but positive communication rates are achievable by using *degenerate quantum error correcting codes* [7, 10, 11] to suppress the input-environment mutual information. The threshold value of p where the capacity goes to zero is unknown. For p close to zero, the best lower bound for $Q(\mathcal{D}_p)$ is the one-shot coherent information, while the best analytical upper bound is at least $O(p)$ higher [13, 14]. Recently, [15] found a numerical upper bound on $Q(\mathcal{D}_p)$ which is very close to the 1-shot coherent information for small p. Meanwhile, the complementary channel for the depolarizing channel for any $p > 0$ is found to always have positive capacity [12], which renders several other techniques to understand the capacity inapplicable [16, 17].

Low-noise channels. In this paper, we consider the quantum capacity of "low-noise quantum channels" that are close to the identity channel, and investigate how close the capacity is to the 1-shot coherent informa-

tion. It has been unclear whether we should expect substantial nonadditivity of coherent information for such channels. On the one hand, all known degenerate codes that boost the quantum capacity above the 1-shot coherent information first encode one logical qubit into a small number of physical qubits, which incurs a significant penalty in rate. This would seem to preclude any benefit in the regime where the 1-shot coherent information is already quite high. On the other hand, we have no effective methods for evaluating the n-letter coherent information for large n, and there may well exist new types of coding strategies that incur no such penalty in the large n regime.

We prove in this paper that to linear order in the noise parameter, the quantum capacity of any low-noise channel *is* its 1-shot coherent information. Consequently, degenerate codes cannot improve the rates of these channels up to the same order. For the special cases of the qubit depolarizing channel, the mixed Pauli channel and their qudit generalizations, we show that the quantum capacity and the 1-shot coherent information agree to even higher order.

Our findings extend to the private capacity $P(\mathcal{N})$ of a quantum channel $\mathcal{N}$ (transmitting classical data of which the environment has little information). Similar to the quantum capacity, the private capacity is an optimized function called the private information optimized on n uses of $\mathcal{N}$ divided by n [5], and the private information is not additive ([18–20] and not well understood, see [2] for details). Since quantum transmission is necessarily private, the private capacity is never smaller than the quantum capacity. Reference [21] exhibits channels with positive private capacity but zero quantum capacity, and characterizes noise that hurts quantum transmission and that can be "shielded" from corrupting private data. In [22], channels are found with almost no quantum capacity but maximum private capacity. Meanwhile, for degradable channels $\mathcal{N}$, the private capacity is again equal to the 1-shot coherent information, $P(\mathcal{N}) = I_c(\mathcal{N})$ [23]. We apply our techniques to show that the private capacity of low-noise channels is also equal to the 1-shot coherent information, and to the quantum capacity to linear order in the noise parameter. Consequently, shielding provides little benefit.

Our results follow closely the approach in [15]. Recall that a quantum channel $\mathcal{N}\colon A \to B$ can be described as an isometry U from the input system A to an output B and an environment E, followed by discarding the environment, $\mathcal{N}(\rho) = \mathrm{tr}_E(U\rho U^\dagger)$. A *complementary channel* $\mathcal{N}^c$ mapping the input to the environment is obtained by discarding the output, $\mathcal{N}^c(\rho) = \mathrm{tr}_B(U\rho U^\dagger)$. A channel $\mathcal{N}$ is called degradable if there is another channel $\mathcal{M}$ (called a degrading map) such that $\mathcal{M}\circ\mathcal{N} = \mathcal{N}^c$. For a general channel, [15] considers minimizing the *diamond norm distance* $\|\mathcal{M}\circ\mathcal{N} - \mathcal{N}^c\|_\diamond$ over all degrading maps $\mathcal{M}$. We call the resulting minimum distance $\mathrm{dg}(\mathcal{N})$ the *degradability parameter of* $\mathcal{N}$. Continuity results, relative to the case as if $\mathcal{N}$ is degradable, can then be obtained similarly to [24]. This new bound in [15] limits the difference between the 1-shot coherent information and the quantum capacity to $O(\eta\log\eta)$ where η is the degradability parameter, and a similar result holds for the private capacity:

Theorem 1 ([15] Theorem 3.3)**.** *If $\mathcal{N}$ is a channel with degradability parameter $\mathrm{dg}(\mathcal{N}) = \eta$, then,*

$$\begin{aligned}|Q(\mathcal{N}) - I_c(\mathcal{N})| &\le \frac{\eta}{2}\log(|E|-1) + \eta\log|E| \\ &\quad + h\Big(\frac{\eta}{2}\Big) + \Big(1+\frac{\eta}{2}\Big)\,h\left(\frac{\eta}{2+\eta}\right) \\ |P(\mathcal{N}) - I_c(\mathcal{N})| &\le \eta\log(|E|-1) + 4\eta\log|E| \\ &\quad + 2\,h\Big(\frac{\eta}{2}\Big) + 4\Big(1+\frac{\eta}{2}\Big)\,h\left(\frac{\eta}{2+\eta}\right),\end{aligned}$$

where $h(x) := -x\log x - (1-x)\log(1-x)$ is the binary entropy function, and $|E|$ is the Choi rank of $\mathcal{N}$ [2].

Note that $\eta\log\eta$ does not have a finite slope at $\eta = 0$ but it goes to zero faster than η^b for any $b < 1$. This new bound is not explicit in general, due to the minimization needed to evaluate η. However, the optimization for η is a semidefinite program, giving good numerical and some analytic access to the quantity, and thus to the resulting capacity bounds.

The primary contribution in this paper is an analytic proof of a surprising fact that, for low-noise channels whose diamond norm distance to being noiseless is ε, the degradability parameter η grows at most as fast as $O(\varepsilon^{1.5})$, rendering the gap $O(\eta\log\eta)$ between the 1-shot coherent information and the quantum or private capacity only sublinear in ε (see Theorem 2). For the qubit depolarizing channel and its various generalizations, we improve the analytic bound of η to $O(\varepsilon^2)$ (see Theorem 4 in this paper and Theorem 19 in [2]). Furthermore, we provide constructive approximate degrading maps and explain why they work well.

In the following, we provide more detailed descriptions and derivations of our results. The diamond norm distance between two quantum channels $\mathcal{N}_{1,2}\colon A \to B$ can be expressed as $\|\mathcal{N}_1 - \mathcal{N}_2\|_\diamond = \max_\rho \|(\mathrm{id}\otimes(\mathcal{N}_1 - \mathcal{N}_2))(\rho)\|_1$ where the maximization is over Hermitian matrices with unit trace norm, $\|\rho\|_1 = 1$, and where id is the identity channel on a reference system isomorphic to the input space of $\mathcal{N}_{1,2}$ (see [2] for details). For a quantum channel $\mathcal{N}\colon A \to B$ with complementary channel $\mathcal{N}\colon A \to E$, the coherent information is defined as $I_c(\mathcal{N}) := \max_\rho\{S(\mathcal{N}(\rho)) - S(\mathcal{N}^c(\rho))\}$, where $S(\sigma) = -\mathrm{tr}\,\sigma\log\sigma$ denotes the von Neumann entropy.

Main results. We first derive an upper bound for the degradability parameter of a general low-noise quantum channel $\mathcal{N}$ satisfying $\|\mathcal{N} - \mathrm{id}\,\|_\diamond \le \varepsilon$ (here, $A = B$). We want a channel $\mathcal{M}$ such that $\mathcal{M}\circ\mathcal{N} \approx \mathcal{N}^c$. Since

$\mathcal{N} \approx \mathrm{id}$, the complementary channel $\mathcal{M} = \mathcal{N}^c$ is a rather good choice.

Theorem 2. *Let* $\|\mathcal{N} - \mathrm{id}\|_\diamond \leq \varepsilon$ *for* $\varepsilon \in [0,2]$. *Then*

$$\|\mathcal{N}^c - \mathcal{N}^c \circ \mathcal{N}\|_\diamond \leq 2\varepsilon^{1.5}.$$

Thus, $\mathrm{dg}(\mathcal{N}) \leq 2\varepsilon^{1.5}$.

Proof. For the quantum channel $\mathcal{N}\colon A \to A$, let R be a system isomorphic to A, and let $\rho = \rho_{RA}$ be a state such that $\|\mathcal{N}^c - \mathcal{N}^c \circ \mathcal{N}\|_\diamond = \|(\mathrm{id} \otimes \mathcal{N}^c)(\rho) - (\mathrm{id} \otimes \mathcal{N}^c \circ \mathcal{N})(\rho)\|_1$. We set $\delta = \rho - (\mathrm{id} \otimes \mathcal{N})(\rho)$, which has a spectral decomposition $\delta = \sum \lambda_i |\psi_i\rangle\langle\psi_i|$. Since $\mathcal{N}$ is trace preserving, $\mathrm{id} \otimes \mathrm{tr}(\delta) = 0$. Also, $\|\mathcal{N} - \mathrm{id}\|_\diamond \leq \varepsilon$ implies $\sum_i |\lambda_i| \leq \varepsilon$. We now use the above facts to compute

$$\begin{aligned}
&\|\mathcal{N}^c - \mathcal{N}^c \circ \mathcal{N}\|_\diamond \\
&\quad = \|\,\mathrm{id} \otimes \mathcal{N}^c(\delta)\|_1 \\
&\quad = \|\,\mathrm{id} \otimes \mathcal{N}^c(\delta) - \mathrm{id} \otimes \mathrm{tr}(\delta)\|_1 \\
&\quad = \|\sum \lambda_i (\mathrm{id} \otimes \mathcal{N}^c - \mathrm{id} \otimes \mathrm{tr})(|\psi_i\rangle\langle\psi_i|)\|_1 \\
&\quad \leq \sum |\lambda_i| \|(\mathrm{id} \otimes \mathcal{N}^c - \mathrm{id} \otimes \mathrm{tr})(|\psi_i\rangle\langle\psi_i|)\|_1 \\
&\quad \leq (\sum |\lambda_i|) \|\mathcal{N}^c - \mathrm{tr}\,\|_\diamond \\
&\quad \leq \varepsilon \|\mathcal{N}^c - \mathrm{tr}\,\|_\diamond.
\end{aligned}$$

Since $\mathrm{tr} = \mathrm{id}^c$, we have $\|\mathcal{N}^c - \mathrm{id}^c\,\|_\diamond \leq 2\sqrt{\varepsilon}$ by the continuity of the Stinespring representation (Corollary 9 in [2]), and this concludes the proof. □

Theorem 2 bounding the degradability parameter of a low-noise channel, together with the approximate degradability bounds in Theorem 1, immediately gives us tight bounds on the channel's quantum and private capacities.

Theorem 3. *Let* $\|\mathcal{N} - \mathrm{id}\|_\diamond \leq \varepsilon$ *for* $\varepsilon \in [0,1]$. *Then*

$$\begin{aligned}
|Q(\mathcal{N}) - I_c(\mathcal{N})| &\leq f_1(2\varepsilon^{3/2}) = O(\varepsilon^{3/2} \log \varepsilon) \\
|P(\mathcal{N}) - I_c(\mathcal{N})| &\leq f_2(2\varepsilon^{3/2}) = O(\varepsilon^{3/2} \log \varepsilon),
\end{aligned}$$

where $f_1(\delta) = \frac{\delta}{2}\log(|E|-1) + h\left(\frac{\delta}{2}\right) + \delta \log |E| + \left(1 + \frac{\delta}{2}\right) h\left(\frac{\delta}{2+\delta}\right)$ *and* $f_2(\delta) = \delta \log(|E|-1) + 2h\left(\frac{\delta}{2}\right) + 4\delta \log |E| + 4\left(1 + \frac{\delta}{2}\right) h\left(\frac{\delta}{2+\delta}\right)$.

We now turn our attention to the depolarizing channel with error $p \in [0,1]$,

$$\mathcal{D}_p(\rho) := (1-p)\rho + \frac{p}{3}(X\rho X + Y\rho Y + Z\rho Z), \quad (1)$$

where X, Y, Z are the usual Pauli matrices. Since $\|\mathcal{D}_p - \mathrm{id}\,\| = 2p$ (see [2]), Theorem 2 tells us that $Q(\mathcal{D}_p)$ and $P(\mathcal{D}_p)$ differ from $I_c(\mathcal{D}_p) = 1 - h(p) - p\log 3$ by no more than $O(p^{1.5}\log p)$. In the following, we improve the bound for this channel to $O(p^2 \log p)$.

We would like to find $\mathcal{M}_p$ such that $\mathcal{D}_p^c \approx \mathcal{M}_p \circ \mathcal{D}_p$. Since $\mathcal{D}_p \approx \mathrm{id}$, our zeroth-order guess is again $\mathcal{M}_p = \mathcal{D}_p^c$. However, since $\mathcal{D}_p$ is slightly *noisier* than id, we will do better by choosing $\mathcal{M}_p$ to be slightly *less* noisy than $\mathcal{D}_p^c$. In particular, we set $\mathcal{M}_p = \mathcal{D}_{s(p)}^c$, the complementary channel to a depolarizing channel with $s(p) = p + ap^2$ for some $a > 0$ (note that $s(p) = p + ap^2 \leq 1$ in the low-noise regime $p \gtrsim 0$, and hence $\mathcal{D}_{s(p)}^c$ is a valid quantum channel). Since $\mathcal{D}_{s(p)}$ is noisier than $\mathcal{D}_p$, the complementary channel $\mathcal{D}_{s(p)}^c$ is less noisy than $\mathcal{D}_p^c$. Using this ansatz and tuning a appropriately leads to the following theorem.

Theorem 4. *Let* $\mathcal{D}_p$ *be a depolarizing channel with error probability* p. *Then,*

$$\|\mathcal{D}_p^c - \mathcal{D}_{p+\frac{8}{3}p^2}^c \circ \mathcal{D}_p\|_\diamond \leq \frac{64}{3}p^2 + O(p^{5/2}).$$

Thus, $\mathrm{dg}(\mathcal{D}_p) \leq \frac{64}{3}p^2 + O(p^{5/2})$.

In fact, the leading constant in Theorem 4 can be improved from $\frac{64}{3} \approx 21.3$ to $\frac{8}{9}(6+\sqrt{2}) \approx 6.6$, which we prove in Theorem 17 in the supplementary material [2].

Proof of Theorem 4. The complementary channel of $\mathcal{D}_p$, which we refer to as the *epolarizing channel* (cf. [12]), can be chosen such that its action on a linear operator ρ is given by $\mathcal{D}_p^c(\rho) =$

$$\begin{pmatrix}
(1-p)\,\mathrm{tr}(\rho) & b(p)\langle X,\rho\rangle & b(p)\langle Y,\rho\rangle & b(p)\langle Z,\rho\rangle \\
b(p)\langle X,\rho\rangle & \frac{p}{3}\,\mathrm{tr}(\rho) & -\frac{ip}{3}\langle Z,\rho\rangle & \frac{ip}{3}\langle Y,\rho\rangle \\
b(p)\langle Y,\rho\rangle & \frac{ip}{3}\langle Z,\rho\rangle & \frac{p}{3}\,\mathrm{tr}(\rho) & -\frac{ip}{3}\langle X,\rho\rangle \\
b(p)\langle Z,\rho\rangle & -\frac{ip}{3}\langle Y,\rho\rangle & \frac{ip}{3}\langle X,\rho\rangle & \frac{p}{3}\,\mathrm{tr}(\rho)
\end{pmatrix}, \quad (2)$$

where $\langle P, Q\rangle := \mathrm{tr}(P^\dagger Q)$ is the Hilbert-Schmidt inner product between operators, and $b(p) := \sqrt{\frac{p(1-p)}{3}}$. We now show that there is a value of a such that $\|\mathcal{D}_p^c - \mathcal{D}_{p+ap^2}^c \circ \mathcal{D}_p\|_\diamond = O(p^2)$. Using (1) and (2), we obtain $\mathcal{D}_{p+ap^2}^c \circ \mathcal{D}_p(\rho) =$

$$\begin{pmatrix}
(1-s)\,\mathrm{tr}(\rho) & b(s)p'\langle X,\rho\rangle & b(s)p'\langle Y,\rho\rangle & b(s)p'\langle Z,\rho\rangle \\
b(s)p'\langle X,\rho\rangle & \frac{s}{3}\,\mathrm{tr}(\rho) & -\frac{is}{3}p'\langle Z,\rho\rangle & \frac{is}{3}p'\langle Y,\rho\rangle \\
b(s)p'\langle Y,\rho\rangle & \frac{is}{3}p'\langle Z,\rho\rangle & \frac{s}{3}\,\mathrm{tr}(\rho) & -\frac{is}{3}p'\langle X,\rho\rangle \\
b(s)p'\langle Z,\rho\rangle & -\frac{is}{3}p'\langle Y,\rho\rangle & \frac{is}{3}p'\langle X,\rho\rangle & \frac{s}{3}\,\mathrm{tr}(\rho)
\end{pmatrix}, \quad (3)$$

where $p' := 1 - \frac{4p}{3}$ and $s = s(p) = p + ap^2$. We set $\Phi = \mathcal{D}_p^c - \mathcal{D}_{p+ap^2}^c \circ \mathcal{D}_p$, whose action on a linear operator ρ is given by the difference between (2) and (3). To upper bound $\|\Phi\|_\diamond$, we apply Lemma 6 of [2], which states that $\|\Phi\|_\diamond \leq 8\|\mathcal{J}(\Phi)\|_{\max}$, where $\mathcal{J}(\Phi) =$

$\sum_{i,j=0}^{1} |i\rangle\langle j| \otimes \Phi(|i\rangle\langle j|)$ is the *Choi matrix* of Φ, and $\|\cdot\|_{\max}$ denotes the maximum absolute value over the entries of a matrix. Due to the block structure of the Choi matrix, $\|\mathcal{J}(\Phi)\|_{\max} = \max_{i,j} \|\Phi(|i\rangle\langle j|)\|_{\max}$. To find this maximum, first note that for any i, j, the quantity $|\langle X, |i\rangle\langle j|\rangle|$ is either 0 or 1, and similarly for $|\langle Y, |i\rangle\langle j|\rangle|$ and $|\langle Z, |i\rangle\langle j|\rangle|$. Therefore, from inspection of the difference between (2) and (3), $\max_{i,j} \|\Phi(|i\rangle\langle j|)\|_{\max}$ is either $s - p = ap^2$, or $\frac{1}{3}|sp' - p| = \frac{1}{3}|a - \frac{4}{3}|p^2 + O(p^3)$, or $b(p) - p'b(s)$. The latter has a Taylor series expansion around $p = 0$ as $(\frac{4}{3\sqrt{3}} - \frac{a}{2\sqrt{3}})p^{3/2} + O(p^{5/2})$, which is $O(p^{5/2})$ if $a = \frac{8}{3}$. With this choice, $\max_{i,j} \|\Phi(|i\rangle\langle j|)\|_{\max} = ap^2 = \frac{8}{3}p^2$ for sufficiently small p. Altogether, $\|\mathcal{J}(\Phi)\|_\diamond \leq 8\|\mathcal{J}(\Phi)\|_{\max} \leq \frac{64}{3}p^2 + O(p^{5/2})$, which completes the proof. □

Theorem 4, together with Theorem 1, immediately gives us new bounds on the quantum and private capacities of the qubit depolarizing channel.

Theorem 5.

$$\begin{aligned} 0 &\leq Q(\mathcal{D}_p) - (1 - h(p) - p\log 3) \\ &\leq -\frac{128}{3}\, p^2 \log p + O(p^2) \\ 0 &\leq P(\mathcal{D}_p) - (1 - h(p) - p\log 3) \\ &\leq -128\, p^2 \log p + O(p^2), \end{aligned}$$

Furthermore, Theorem 17 in the supplementary material [2] reduces the constants of the RHS by more than a factor of 3. This is the first such bound known for the low-noise depolarizing channel.

Extensions and implications. Our key new finding is that channels within ε of perfect are also exceptionally close to degradable, with degradability parameter of $O(\varepsilon^{1.5})$ in general and $O(p^2)$ for the qubit depolarizing channel $\mathcal{D}_p$. It remains open if the $O(\varepsilon^2)$ degradability parameter holds in general, but our proof techniques readily extend to some important classes of quantum channels. In particular, we establish explicit quadratic upper bounds for the degradability parameter for the *Pauli channels*, which generalizes the depolarizing channel to arbitrary probabilities of having the four Pauli errors (see Theorem 19 in [2] for details). Furthermore, similar results hold for higher dimensional generalizations of the Pauli channels.

An important instance of the Pauli channels is the XZ-channel (often called the BB84-channel in a quantum key distribution context)

$$\begin{aligned} \mathcal{N}_{p,q}^{XZ}(\rho) \coloneqq (1-p)(1-q)\rho + p(1-q)X\rho X \\ + pqY\rho Y + (1-p)qZ\rho Z, \end{aligned}$$

which implements independently an X-dephasing with probability p, and a Z-dephasing with probability q. For our discussion, we set $p = q$ and denote the resulting XZ-channel by $\mathcal{C}_p \coloneqq \mathcal{N}_{p,p}^{XZ}$, which has coherent information $I_c(\mathcal{C}_p) = 1 - 2h(p)$. Using similar methods as in the proof of Theorem 4, we show $\mathrm{dg}(\mathcal{C}_p) \leq 64p^2 + O(p^3)$ in [2]. Moreover, similar to the depolarizing channel the coefficient of p^2 can be improved from 64 to 16, which we show in Theorem 21 in [2].

The nonadditivity of coherent information for a general channel implies that degenerate codes are sometimes needed to achieve the quantum capacity [7, 8, 10, 11, 20, 25], but little is known about these codes despite 20 years of research. We showed that the coherent information is essentially the quantum capacity for low-noise channels. Therefore, we have also arrived at a refreshing result that using random block codes on the typical subspace of the optimal input (for the 1-shot coherent information) essentially achieves the capacity.

Likewise, our findings have implications in quantum cryptography. In quantum key distribution, quantum states are transmitted through well-characterized noisy quantum channels that are subject to further adversarial attacks. Parameter estimation is used to determine the effective channel (see for example [26]) and the optimal key rate of one-way quantum key distribution is the private capacity of the effective channel. These effective channels typically have low noise (e.g., $1 - 2\%$ in [27]), and our results imply near-optimality of the simple (classical) error correction and privacy amplification procedures that approach the one-shot coherent information of the effective channel. In particular, for the XZ-channel with bit-flip probability p, the optimal key rate is $1 - 2h(p) + O(p^2 \log p)$.

Finally, our results can be extended to *generalized low-noise* channels $\mathcal{N}$, for which there exists another channel $\mathcal{M}$ such that $\|\mathcal{M} \circ \mathcal{N} - I\|_\diamond \leq \epsilon$. For example, this includes all channels that are close to isometric channels. For a generalized low-noise channel, we have by Theorem 2 that

$$\|(\mathcal{M} \circ \mathcal{N})^c - (\mathcal{M} \circ \mathcal{N})^c \circ (\mathcal{M} \circ \mathcal{N})\|_\diamond \leq 2\epsilon^{3/2}. \quad (4)$$

Furthermore, up to an isometry,

$$(\mathcal{M} \circ \mathcal{N})^c(\rho) = (\mathcal{M}^c \otimes I_{E_1})(U_{\mathcal{N}} \rho U_{\mathcal{N}}^\dagger),$$

where $U_{\mathcal{N}} : A \to BE_1$ is an isometric extension of $\mathcal{N}$ and $\mathcal{M}^c : B \to E_2$, so that $\mathrm{tr}_{E_2}(\mathcal{M} \circ \mathcal{N})^c(\rho) = \mathcal{N}^c(\rho)$. Equation (4) therefore implies

$$\|\mathcal{N}^c - \mathrm{tr}_{E_2}(\mathcal{M} \circ \mathcal{N})^c \circ (\mathcal{M} \circ \mathcal{N})\|_\diamond \leq 2\epsilon^{3/2},$$

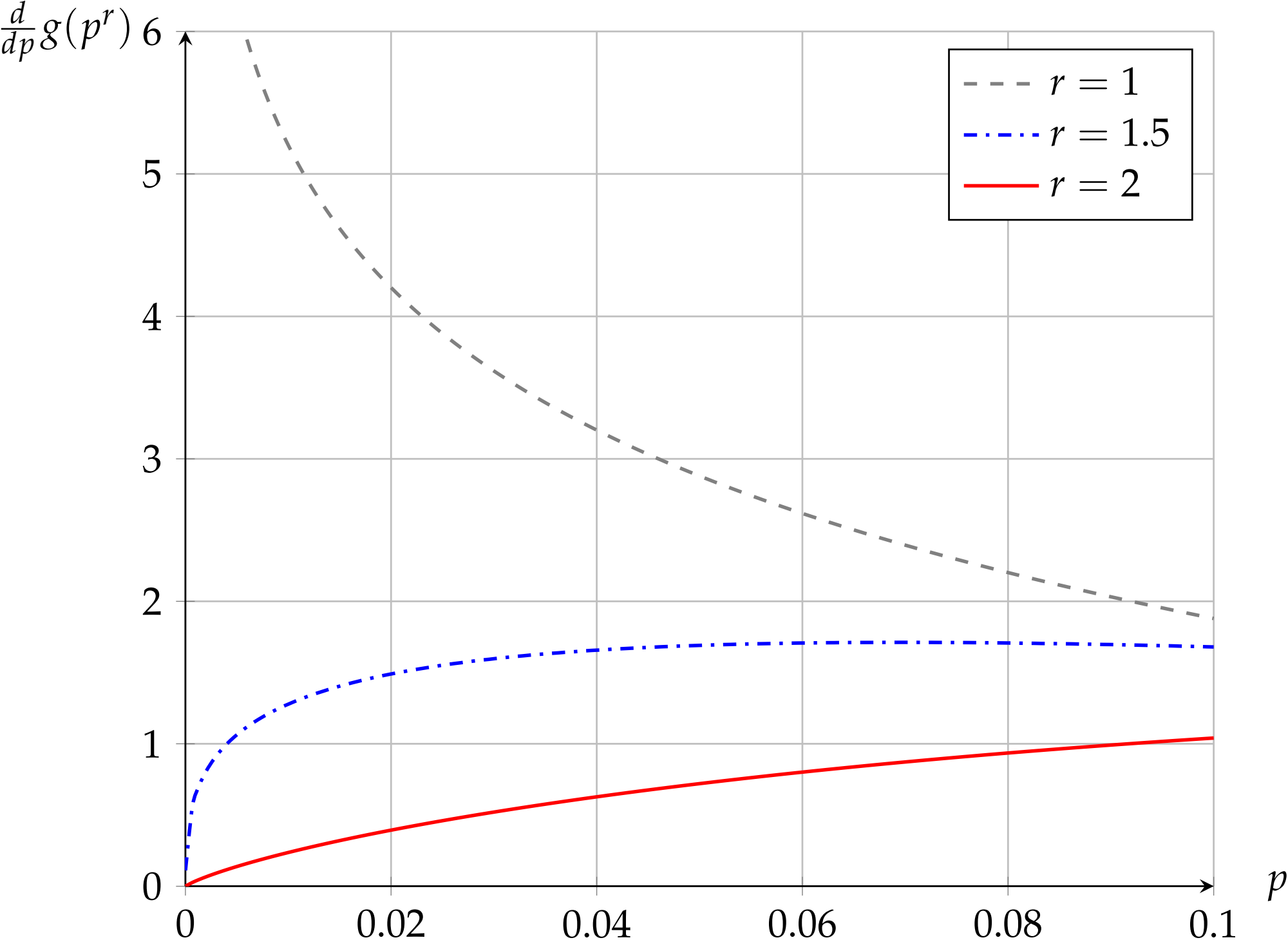
so that letting $\mathcal{D} = \mathrm{tr}_{E_2}(\mathcal{M} \circ \mathcal{N})^c \circ \mathcal{M}$ we have $\|\mathcal{N}^c - \mathcal{D} \circ \mathcal{N}\|_\diamond \leq 2\epsilon^{3/2}$ and $\mathcal{N}$ is approximately degradable with degradability parameter $2\epsilon^{3/2}$. We thus find that the same bounds as in Theorem 3 apply in the case of a generalized low-noise channel $\mathcal{N}$.

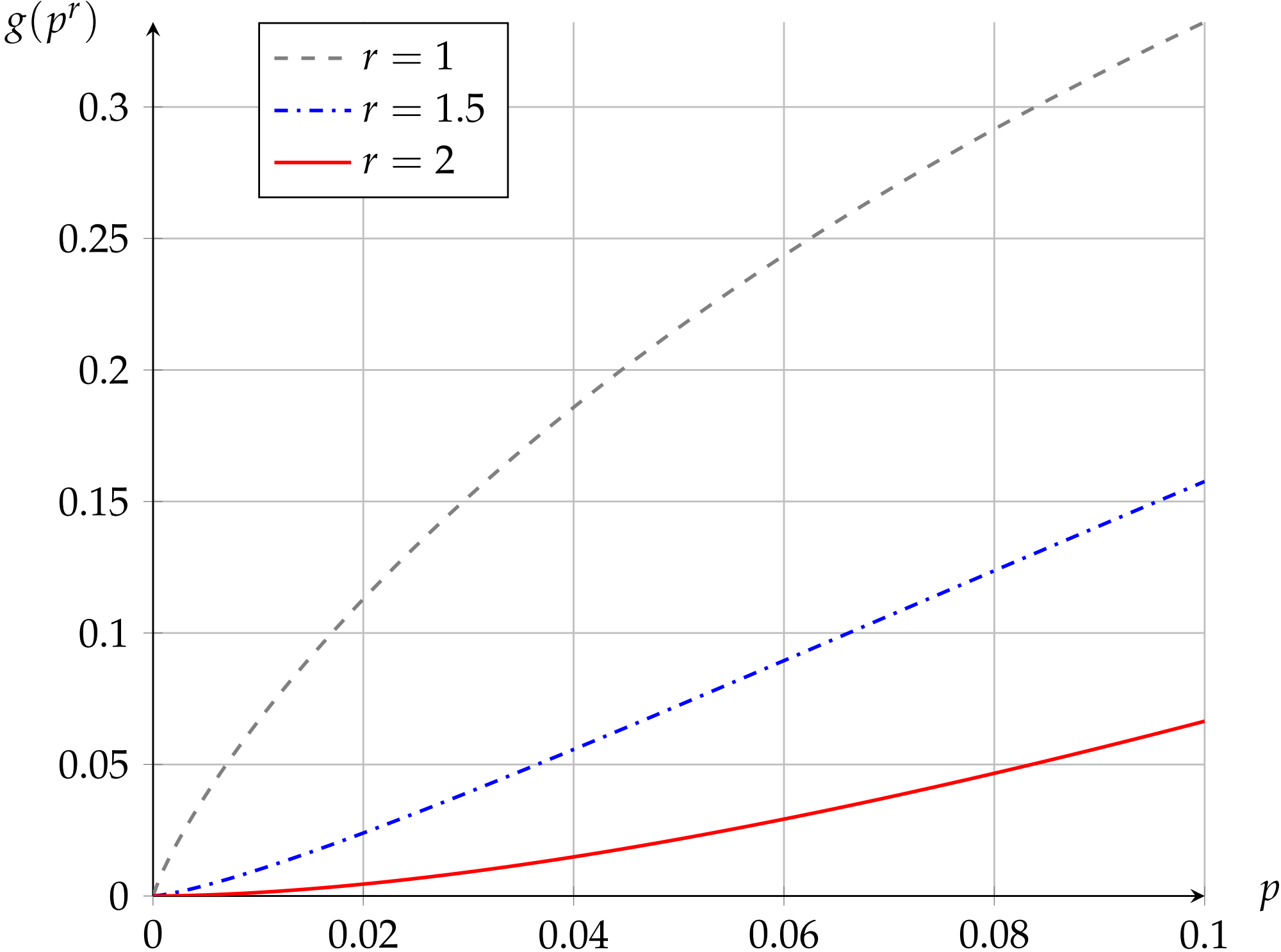
Acknowledgements. We thank Charles Bennett, Ke Li, John Smolin and John Watrous for inspiring discussions, and Mark M. Wilde for helpful feedback. DL is further supported by NSERC and CIFAR, and FL and GS by the National Science Foundation under Grant Number 1125844.

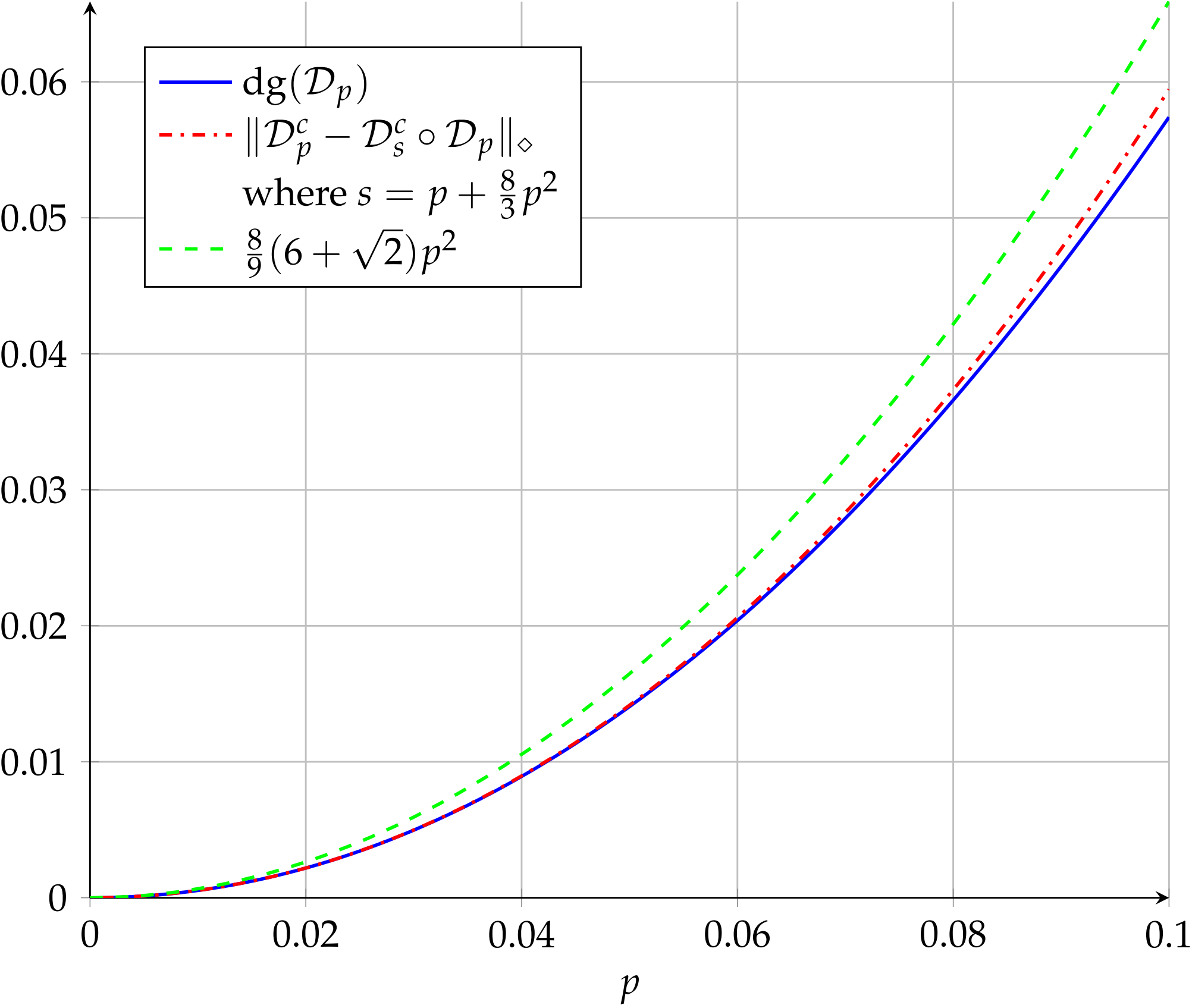
---

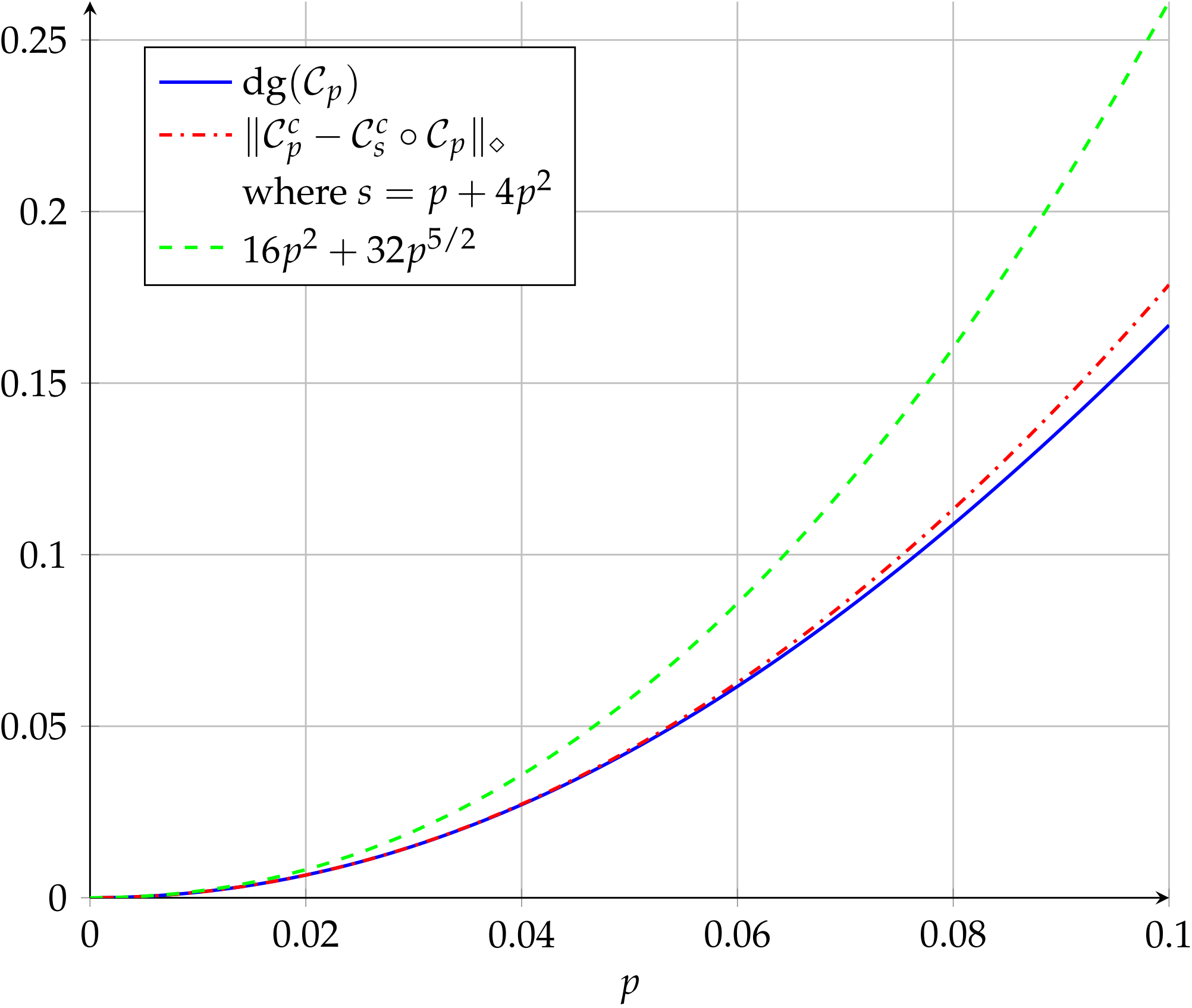
* Electronic address: felix.leditzky@jila.colorado.edu
† Electronic address: wcleung@iqc.ca
‡ Electronic address: gsbsmith@gmail.com

[1] C. Shannon, The Bell System Technical Journal **27**, 379 (1948).
[2] *See supplementary material for a full account of the mathematical details of the main text, which includes Refs. [28–37].*
[3] S. Lloyd, Physical Review A **55**, 1613 (1997), arXiv:quant-ph/9604015.
[4] P. Shor, *The quantum channel capacity and coherent information*, Lecture notes, MSRI Workshop on Quantum Computation (2002), URL http://www.msri.org/realvideo/ln/msri/2002/quantumcrypto/shor/1.
[5] I. Devetak, IEEE Transactions on Information Theory **51**, 44 (2005), arXiv:quant-ph/0304127.
[6] I. Devetak and P. W. Shor, Communications in Mathematical Physics **256**, 287 (2005), arXiv:quant-ph/0311131.
[7] D. P. DiVincenzo, P. W. Shor, and J. A. Smolin, Physical Review A **57**, 830 (1998), arXiv:quant-ph/9706061.
[8] T. Cubitt, D. Elkouss, W. Matthews, M. Ozols, D. Pérez-García, and S. Strelchuk, Nature Communications **6** (2015), arXiv:1408.5115 [quant-ph].
[9] D. Bruß, D. P. DiVincenzo, A. Ekert, C. A. Fuchs, C. Macchiavello, and J. A. Smolin, Physical Review A **57**, 2368 (1998), arXiv:quant-ph/9705038.
[10] G. Smith and J. A. Smolin, Physical Review Letters **98**, 030501 (2007), arXiv:quant-ph/0604107.
[11] J. Fern and K. B. Whaley, Physical Review A **78**, 062335 (2008), arXiv:0708.1597 [quant-ph].
[12] D. Leung and J. Watrous, Quantum **1**, 28 (2017), arXiv:1510.01366 [quant-ph].
[13] G. Smith and J. A. Smolin, in *2008 IEEE Information Theory Workshop* (2008), pp. 368–372, arXiv:0712.2471 [quant-ph].
[14] Y. Ouyang, Quantum Information & Computation **14**, 917 (2014), arXiv:1106.2337 [quant-ph].
[15] D. Sutter, V. B. Scholz, A. Winter, and R. Renner, IEEE Transactions on Information Theory **63**, 7832 (2017), arXiv:1412.0980 [quant-ph].
[16] S. Watanabe, Physical Review A **85**, 012326 (2012), arXiv:1110.5746 [quant-ph].
[17] A. Cross, K. Li, and G. Smith, Physical Review Letters **118**, 040501 (2017), arXiv:1601.05434 [quant-ph].
[18] G. Smith, J. M. Renes, and J. A. Smolin, Physical Review Letters **100**, 170502 (2008), arXiv:quant-ph/0607018.
[19] K. Li, A. Winter, X. Zou, and G. Guo, Physical Review Letters **103**, 120501 (2009), arXiv:0903.4308 [quant-ph].
[20] G. Smith and J. A. Smolin, Physical Review Letters **103**, 120503 (2009), arXiv:0904.4050 [quant-ph].
[21] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, Physical Review Letters **94**, 160502 (2005), arXiv:quant-ph/0309110.
[22] D. Leung, K. Li, G. Smith, and J. A. Smolin, Physical Review Letters **113**, 030502 (2014), arXiv:1312.4989 [quant-ph].
[23] G. Smith, Physical Review A **78**, 022306 (2008), arXiv:0705.3838 [quant-ph].
[24] D. Leung and G. Smith, Communications in Mathematical Physics **292**, 201 (2009), arXiv:0810.4931 [quant-ph].
[25] G. Smith and J. Yard, Science **321**, 1812 (2008), arXiv:0807.4935 [quant-ph].
[26] R. Renner, N. Gisin, and B. Kraus, Physical Review A **72**, 012332 (2005), arXiv:quant-ph/0502064.
[27] D. Stucki, N. Walenta, F. Vannel, R. T. Thew, N. Gisin, H. Zbinden, S. Gray, C. Towery, and S. Ten, New Journal of Physics **11**, 075003 (2009), arXiv:0903.3907 [quant-ph].
[28] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information* (Cambridge University Press, Cambridge, 2000).
[29] J. Watrous, *Theory of Quantum Information* (Cambridge University Press (to be published), 2017), https://cs.uwaterloo.ca/ watrous/TQI/.
[30] W. F. Stinespring, Proceedings of the American Mathematical Society **6**, 211 (1955).
[31] C. W. Helstrom, Journal of Statistical Physics **1**, 231 (1969).
[32] J. Watrous, arXiv preprint (2009), arXiv:0901.4709 [quant-ph].
[33] J. Watrous, arXiv preprint (2012), arXiv:1207.5726 [quant-ph].
[34] D. Kretschmann, D. Schlingemann, and R. F. Werner, IEEE Transactions on Information Theory **54**, 1708 (2008), arXiv:quant-ph/0605009.
[35] M. Fekete, Mathematische Zeitschrift **17**, 228 (1923).
[36] J. Yard, P. Hayden, and I. Devetak, IEEE Transactions on Information Theory **54**, 3091 (2008), arXiv:quant-ph/0501045.
[37] F. Leditzky, E. Kaur, N. Datta, and M. M. Wilde, Physical Review A **97**, 012332 (2018), arXiv:1709.01111 [quant-ph].


$\frac{d}{dp}g(p^r)$
6
5
4
3
2
1
0
0
0.02
0.04
0.06
0.08
0.1
p
$r = 1$
$r = 1.5$
$r = 2$


$g(p^r)$
$r = 1$
$r = 1.5$
$r = 2$
0.3
0.25
0.2
0.15
0.1
0.05
0
0
0.02
0.04
0.06
0.08
0.1
p


$\mathrm{dg}(\mathcal{D}_p)$
$\|\mathcal{D}_p^c - \mathcal{D}_s^c \circ \mathcal{D}_p\|_\diamond$
where $s = p + \frac{8}{3}p^2$
$\frac{8}{9}(6+\sqrt{2})p^2$
0.06
0.05
0.04
0.03
0.02
0.01
0
0
0.02
0.04
0.06
0.08
0.1
p


$\mathrm{dg}(\mathcal{C}_p)$
$\|\mathcal{C}_p^c - \mathcal{C}_s^c \circ \mathcal{C}_p\|_\diamond$
where $s = p + 4p^2$
$16p^2 + 32p^{5/2}$
0.25
0.2
0.15
0.1
0.05
0
0
0.02
0.04
0.06
0.08
0.1
p