1 Accuracy threshold for concatenated DD

Next, we turn to the situation of CDD. We can certainly discuss the break-even point for a particular CDD_n sequence, treating it simply as a "flattened" sequence of pulses, and ask when the error phase after CDD_n is smaller than ϕ_{SB} as was done for PDD in the previous section. However, what is more interesting is the question of the so-called accuracy threshold for CDD. As described in the Preliminaries section, CDD tries to achieve better and better noise removal by concatenating the same basic DD scheme—PDD for our discussion here—to higher and higher levels. Here, we are interested in understanding how the noise removal improves as the DD scheme scales up. In fault-tolerant quantum computing, there is the concept of a fault tolerance noise threshold, a strength of the noise below which scaling up the QEC code leads to improved protection against noise, and hence more accurate quantum computation. This threshold is referred to as the accuracy threshold. Here, we ask the analogous question of CDD, whether there is a condition on the noise parameter of the problem such that increasing the CDD level always leads to improved noise removal capabilities. An accuracy threshold for CDD exists if the parameters governing the noise of the situation satisfying a condition $\eta < \eta_0$, such that

$$\Phi_{SB,n+1} < \Phi_{SB,n} \qquad \forall n = 1, 2, \dots. \tag{1}$$

Below, we examine this for both the ideal CDD case where pulses are perfect, and the noisy CDD case where imperfections in the pulses are allowed.

1.1 Ideal case

If a noise threshold indeed exist, then ideal case, where the noise is non-existent, would definitely apply. We first consider CDD with perfect pulses. Khodjasteh and Lidar propose in their paper to truncate to th second order term gave in their paper an estimation of the CDD_n generator as:

$$
\begin{aligned}
\Omega_{(n)} \approx\ & \sigma_0 \otimes 4^n B_0 \\
& + \sigma_1 \otimes (-\mathrm{i})^n 2^{n(n+1)} \operatorname{ad}_{B_0}^n (B_1) \\
& + \sigma_2 \otimes (-\mathrm{i})^n 2^{n^2} \operatorname{ad}_{B_0}^{n-1} ([B_0, B_2] - \mathrm{i}\{B_1, B_3\}).
\end{aligned}
\tag{2}
$$

It is shown in the appendix A that the above expression indeed captures the leading order behavior of $\Omega_{(n)}$. Different from the leading order term in the Magnus series, the leading order "behavior" is manifested separately for the pure bath part and the interaction part. As mentioned in the Preliminaries section, CDD_n achieves nth-order decoupling. From the leading order expression, we see it is indeed so as the interaction part is of $(n+1)$th order. From Eq. (2), we can estimate $\Phi_{\text{B},n} \simeq 4^n \phi_\text{B}$ and derive an upper bound for for the error phase:

$$
\Phi_{\text{SB},n} \lesssim 2^{(n+1)^2} \phi_\text{B}^n\, \phi_\text{SB}\, f_n\!\left(\frac{\phi_\text{SB}}{\phi_\text{B}}\right),
\tag{3}
$$

where we define the piece-wise function:

$$
f_n(x) =
\begin{cases}
1, & x \le \sqrt{4^n - 1}, \\[2mm]
\dfrac{1}{2^n} \sqrt{\left(\dfrac{4^n - 1}{2x} + \dfrac{x}{2}\right)^2 + 1}, & x > \sqrt{4^n - 1}.
\end{cases}
\tag{4}
$$

to remind the reader, $\phi_\text{B} \equiv \phi_\text{B}(\tau_0) \equiv \tau_0 \|H_\text{B}\|$, and $\phi_\text{SB} \equiv \phi_\text{SB}(\tau_0) \equiv \tau_0 \|H_\text{SB}\|$, noting the dependence of both quantities on τ_0, the time interval between two consecutive pulses in the CDD sequence. (To reconsider: We need both an upper and lower bound for $\Phi_{\text{SB},n}$, which we have from the estimator [Eq. (11)]. For use in Eq. (1), we should have the lower bound on the RHS of that condition, and upper bound expression on the LHS, for a sufficient condition.) As mentioned in the Preliminaries section, CDD_n

achieves nth-order decoupling. This can be understood from the error phase expression: $\Phi_{\mathrm{SB},n}$ after CDD_n is of order ϕ^{n+1}, for $\phi \sim \phi_{\mathrm{B}}, \phi_{\mathrm{SB}}$. That ϕ_{B} enters the error phase should again be of no surprise—as in the PDD case, ϕ_{B} determines how quickly the noise seen by the system evolves, and hence affects the efficacy of DD which does a good elimination of the noise only when the noise remains nearly unchanged for the full DD sequence.

What is perhaps more surprising is the appearance of factors of 2^n and 2^{n^2} in $\Phi_{\mathrm{SB},n}$, a new feature in the CDD case (Do these factors also appear in the Khodjasteh-Lidar analysis? Please check and put in a sentence about this.). For fixed $\tau_0 \equiv \tau$ as n increases (e.g., if τ is the experimental limit for the switch time between consecutive pulses), such factors mean that the error phase eventually increases for large enough n: The exponentially decreasing ϕ^n factor is eventually overcome by the super-exponentially increasing 2^{n^2} factor. This means there is no accuracy threshold, i.e., there is no level of noise that is weak enough such that $\Phi_{\mathrm{SB},n+1} \leq \Phi_{\mathrm{SB},n}$ for all n. Instead, there is an maximal useful level of concatenation level, beyond which further concatenation actually increases the noise seen by the system. Fig. 1 plots this situation of fixed τ_0, for increasing concatenation level n. We observe the initial decrease of $\Phi_{\mathrm{SB},n}$ as n increases, but this turns around eventually.

(To reconsider with my lower/upper bound comment above.) For higher concatenation level to make sense, increasing concatenation level should decrease the error phase. To derive a sufficient condition of it, we examine the recursive relation between concatenation levels, which suggests

$$
\begin{aligned}
B_{n+1,0} &= 4B_{n,0} \\
B_{n+1,1} &= -4\mathrm{i}[B_{n,0}, B_{n,1}] \\
B_{n+1,2} &= -2\mathrm{i}[B_{n,0}, B_{n,1}] \\
B_{n+1,3} &= 0
\end{aligned}
\tag{5}
$$

For both B_1 and B_2 to decrease in size, it suffice to require the map $-4\mathrm{i}[B_{n,0}, \cdot]$ a contraction map. By requiring its norm to be smaller than 1, we obtain the bound for the maximal concatenation level:

$$
n_{\max} \leq \lceil -\log_4(\phi_{\mathrm{B}}) - 3/2 \rceil,
\tag{6}
$$

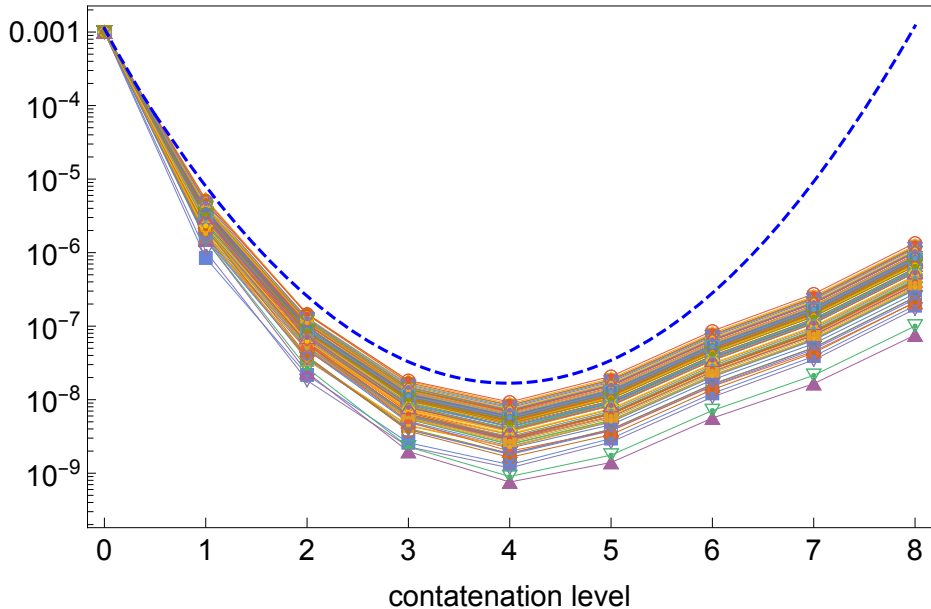where $\lceil \cdot \rceil$ is the ceiling operator.

Figure 1: The error phase $\phi_{\mathrm{SB},n}$ as a function of concatenation level n. The thin color lines are obtained from numerically simulating CDD_n with randomly generated bare Hamiltonian operators satisfying $\phi_{\mathrm{B}} = 1 \times 10^{-3}$ and $\phi_{\mathrm{SB}} = 2 \times 10^{-3}$. The blue dashed line is our theoretical upper bound. For this configuration, we set the maximal concatenation level at 4.

Note that the existence of a maximal useful concatenation level does not technically contradict the statement that higher-order decoupling is achieved by higher level CDD. Using decoupling order to quantify the degree of noise removal requires, in the first place, the convergence of the Magnus series, so that the $(n+1)$th- and higher-order terms in the series are a small correction to the first n terms. However, if the DD scheme is designed such that the total time for the sequence grows exponentially with n, as is the case for CDD if τ_0 is fixed as n increases, eventually, we exceed the convergence criterion and the decoupling order stops being a reasonable indicator of successful noise removal.

We can, however, analyze a different situation: to have fixed $\tau_n \equiv T$, so that the CDD sequence takes the same amount of time, regardless of n. This requires $\tau_0 = \frac{T}{4^n}$ for each n, so that the pulses are applied at shorter and shorter time intervals as the concatenation level increases. This can be the practical choice if one is doing computation where computational gates, which can only be applied at the end of a complete DD sequence in order to not interfere with the noise averaging process, have to be applied at a particular clock rate. Increasing n to increase the noise removal capabilities must not increase the total time taken for the DD sequence. This requires, of course, the ability to do faster and faster pulse switching.

In this case, it is more illuminating to make use of $\tau_0 \|H_{\mathrm{B}}\|$ and $\tau_0 \|H_{\mathrm{SB}}\|$, in place of ϕ_{B} and ϕ_{SB}, in our estimate of the CDD error phases, so that that the n dependence in τ_0 is explicit. (I suggest we analyze this case as well, to see if an accuracy threshold exists. Jiaan - could you take a quick look at this?) Eventually, of course, this becomes limited by the technological—not fundamental—constraint of how fast we can switch between pulses, and thus how short τ_0 can be in an experiment.

1.2 Noisy pulses

(To add.) This has significant implication for the concatenate DD scheme as well, as the first order coupling would persist throughout all concatenation levels.

A   Analyzing CDD_n

It was suggested by the authors of CDD that the decoupling order for CDD_n is n [1]. This statement is correct, but their proof is logically flawed. Given its importance, here we explain why the original proof is insufficient and provide a rigorous proof for the decoupling order.

For the generic CDD_n, it is infeasible to flatten out the gates and calculate the combined Magnus series directly. Instead it is easier to study the iteration map

$$\Omega_{(n)} = \Omega_{\text{PDD}}(\Omega_{(n-1)}) = \sum_{m=1}^{\infty} \Omega_{\text{PDD}}^{(m)}(\Omega_{(n-1)}), \tag{7}$$

where the update map Ω_{PDD} is defined in Eq. (??). Backtracking the map from level n to level 0, we have $\Omega_{(n)} = \Omega_{\text{PDD}} \circ \cdots \circ \Omega_{\text{PDD}}(\Omega) \equiv (\Omega_{\text{PDD}})^n(\Omega)$. In the end, the effective Hamiltonian $\Omega_{(n)}$ can be expanded as:

$$\Omega_{(n)} = \sum_{m=1}^{\infty} \Omega_{(n)}^{(m)}(\Omega), \tag{8}$$

where $\Omega_{(n)}^{(m)}$ is the mth order term for $\Omega_{(n)}$ in Ω such that $\Omega_{(n)}^{(m)} \sim \|\Omega\|^m$, with the $n = 1$ case corresponds to PDD. Notice that this series is different from the series in Eq. (7), where the series is in $\Omega_{(n-1)}$ instead of Ω.

To estimate $\Omega_{(n)}$, it suffice to find an analytically solvable estimator $\widehat{\Omega}_{(n)} \approx \Omega_{(n)}$ which reflects leading order behavior of the full series. The leading order behavior is defined separately for the pure-bath part and the system-bath coupling part, as they can be of different orders. In other words, a faithful estimator should contain the first nonzero pure-bath part and the first nonzero system-bath part of the series (8). Formally, we have an estimator-error pair $(\widehat{\Omega}_{(n)}, \delta\Omega_{(n)})$ such that

$$\Omega_{(n)} = \widehat{\Omega}_{(n)} + \delta\Omega_{(n)}. \tag{9}$$

By definition, the error term must be of higher order than the estimator. To quantify this statement, we take the two-component norm for both $\delta\Omega_{(n)}$ and $\widehat{\Omega}_{(n)}$, and demanding

$$\||\delta\Omega_{(n)}|\| \equiv \begin{pmatrix} \delta\alpha_{(n)} \\ \delta\beta_{(n)} \end{pmatrix} \ll \||\widehat{\Omega}_{(n)}|\| \equiv \begin{pmatrix} \widehat{\alpha}_{(n)} \\ \widehat{\beta}_{(n)} \end{pmatrix}, \tag{10}$$

6

where the comparison is implied for both components and should be understood by comparing the leading powers of the polynomials in α and β. In essence, a "smaller" polynomial should have a larger leading power and thus is of higher order smallness. We refer readers to Appendix **??** for more rigorous descriptions on ordering polynomials by their leading powers. Furthermore, after taking the two-component norm and applying triangular inequality for Eq. (9), we find,

$$\widehat{\beta}_{(n)} - \delta\beta_{(n)} \le \beta_{(n)} \le \widehat{\beta}_{(n)} + \delta\beta_{(n)}. \tag{11}$$

According to Eq. (**??**), the nth-order decoupling condition requires $\beta_{(n)}$ to be an $(n+1)$th order polynomial in α and β. It now suffices to show that ① $\widehat{\beta}_{(n)}$ is an $(n+1)$th order polynomial and ② the estimator is faithful.

The original CDD paper (Ref.[1]) constructed an estimator by keeping the first two Magnus terms for each iteration step. In our notation:

$$\widehat{\Omega}_{(n)} \equiv (\Omega_{\mathsf{PDD}}^{(1)} + \Omega_{\mathsf{PDD}}^{(2)})(\widehat{\Omega}_{(n-1)}) = \left(\Omega_{\mathsf{PDD}}^{(1)} + \Omega_{\mathsf{PDD}}^{(2)}\right)^n(\Omega). \tag{12}$$

With $\Omega_{\mathsf{PDD}}^{(1)}, \Omega_{\mathsf{PDD}}^{(2)}$ explicitly defined in Eq. (**??**) and Eq. (**??**), and the starting point $\widehat{\Omega}_{(0)} \equiv \Omega$, the estimator can shown to be specified by the formula:

$$
\begin{aligned}
\widehat{\Omega}_{(n)} = {} & \sigma_0 \otimes 4^n B_0 \\
& + \sigma_1 \otimes (-\mathrm{i})^n 2^{n(n+1)} \, \mathrm{ad}_{B_0}^n(B_1) \\
& + \sigma_2 \otimes (-\mathrm{i})^n 2^{n^2} \, \mathrm{ad}_{B_0}^{n-1}\big([B_0, B_2] - \mathrm{i}\{B_1, B_3\} \big).
\end{aligned}
\tag{13}
$$

After taking the two-component norm of the estimator, we have,

$$\begin{pmatrix} \widehat{\alpha}_{(n)} \\ \widehat{\beta}_{(n)} \end{pmatrix} \simeq \begin{pmatrix} \alpha \\ \alpha^n \beta + \alpha^{n-1}\beta^2 \end{pmatrix}, \tag{14}$$

where $\simeq$ represents equal the leading powers as described in Appendix **??**. This verifies that $\widehat{\beta}_{(n)}$ is a an $(n+1)$th order polynomial.

To complete the theory, it now remains to show that $\widehat{\Omega}_{(n)}$ is indeed faithful. For simplicity, we make the recognition $\alpha \simeq \beta \simeq \tau_0$. This change does not affect the criteria for nth order decoupling as the polynomial

order is unchanged. To show $\delta\alpha_{(n)}$ and $\delta\beta_{(n)}$ are of higher order in τ_0 compared to $\widehat{\alpha}_{(n)}$ and $\widehat{\beta}_{(n)}$, we claim that:

$$\begin{pmatrix} \delta\alpha_{(n)} \\ \delta\beta_{(n)} \end{pmatrix} \lesssim \begin{pmatrix} \tau_0^3 \\ \tau_0^{n+2} \end{pmatrix}. \tag{15}$$

We prove this assertion by mathematical induction. At $n = 1$, i.e. the PDD sequence, the error comes solely from the Magnus series truncation. Taking the two-component norm and applying triangular inequality:

$$|||\delta\Omega_{(1)}||| \leq \sum_{m=3}^{\infty} |||\Omega_{(1)}^{(m)}||| \simeq |||\Omega_{(1)}^{(3)}||| \lesssim \begin{pmatrix} \tau_0^3 \\ \tau_0^3 \end{pmatrix}. \tag{16}$$

By definition, the higher level error term $\delta\Omega_{(n+1)}$ can be expressed with the lower level functions as:

$$\begin{aligned} \delta\Omega_{(n+1)} &= \Omega_{\mathsf{PDD}}(\Omega_{(n)}) - (\Omega_{\mathsf{PDD}}^{(1)} + \Omega_{\mathsf{PDD}}^{(2)})(\widehat{\Omega}_{(n)}) \\ &= \Omega_{\mathsf{PDD}}^{(1)}(\delta\Omega_{(n)}) + \sum_{m=2}^{\infty} \Omega_{\mathsf{PDD}}^{(m)}(\Omega_{(n)}) - \Omega_{\mathsf{PDD}}^{(2)}(\widehat{\Omega}_{(n)}), \quad (17) \end{aligned}$$

where we have used the fact that $\Omega_{\mathsf{PDD}}^{(1)}$ is a linear map. Assuming Eq. (15) holds for level n, we need to show $|||\delta\Omega_{(n+1)}||| \lesssim (\tau_0^3, \tau_0^{n+3})$. The argument in the original CDD paper by Khodjasteh and Lidar for truncating the Magnus series have only showed that the higher order Magnus term is smaller by proving

$$|||\sum_{m=3}^{\infty} \Omega_{\mathsf{PDD}}^{(m)}(\widehat{\Omega}_{(n)})||| \ll |||\widehat{\Omega}_{(n+1)}|||.$$

If we take $\delta\Omega_{(n)} = 0$, Eq. (17) would identity the l.h.s. of above as $|||\delta\Omega_{n+1}|||$, then this inequality would indeed indicate that the error term is of higher order. However Eq. (17) suggests that this argument is not sufficient: as the estimate error $\delta\Omega_{(n+1)}$ comes not only from the truncation error of the same level, but also accumulates from the lower level error $\delta\Omega_{(n)}$. To properly bound the size of the error term, we take the two-component norm on both sides of Eq. (17):

$$\begin{aligned} |||\delta\Omega_{(n+1)}||| &\lesssim |||\Omega_{\mathsf{PDD}}^{(1)}(\delta\Omega_{(n)})||| + |||\Omega_{\mathsf{PDD}}^{(3)}(\Omega_{(n)})||| \\ &\quad + |||\Omega_{\mathsf{PDD}}^{(2)}(\widehat{\Omega}_{(n)} + \delta\Omega_{(n)}) - \Omega_{\mathsf{PDD}}^{(2)}(\widehat{\Omega}_{(n)})|||, \end{aligned} \tag{18}$$

where the infinite sum of Magnus terms higher than third order is lead by the non-vanishing third order. We now examine the size of these terms. The first term in (18) is easy,

$$\||\Omega_{\mathsf{PDD}}^{(1)}(\delta\Omega_{(n)})\|| = \begin{pmatrix} 4\delta\alpha_{(n)} \\ 0 \end{pmatrix} \simeq \begin{pmatrix} \tau_0^3 \\ 0 \end{pmatrix}.$$

To bound the size of the second term, we make two observations for the PDD Magnus series greater than the third order:

1. At least one $B_{i\neq0}$ is involved in the coupling part.

2. At least two $B_{i\neq0}$ is involved in pure bath part.

Therefore, the second term can be bounded by

$$\||\Omega_{\mathsf{PDD}}^{(3)}(\Omega_{(n)})\|| \lesssim \begin{pmatrix} \beta_{(n)}^2(\alpha_{(n)} + \beta_{(n)}) \\ \beta_{(n)}(\alpha_{(n)} + \beta_{(n)})^2 \end{pmatrix} \simeq \begin{pmatrix} \tau_0^{2n+2} \\ \tau_0^{n+3} \end{pmatrix},$$

where we have used $\alpha_{(n)} \lesssim \widehat{\alpha}_{(n)}$ and $\beta_{(n)} \lesssim \widehat{\beta}_{(n)}$ according to the induction hypothesis. This can be alternatively shown by explicitly calculating the third order Magnus term using Eq. (**??**). To bound the third term in (18), we can calculate the difference $\Omega_{\mathsf{PDD}}^{(2)}(\widehat{\Omega} + \delta\Omega) - \Omega_{\mathsf{PDD}}^{(2)}(\widehat{\Omega})$ using the expression for $\Omega_{\mathsf{PDD}}^{(2)}$ from Eq. (**??**). In the end,

$$\||\Omega_{\mathsf{PDD}}^{(2)}(\widehat{\Omega} + \delta\Omega) - \Omega_{\mathsf{PDD}}^{(2)}(\widehat{\Omega})\|| \lesssim \begin{pmatrix} 0 \\ \delta\alpha_{(n)}\delta\beta_{(n)} \end{pmatrix}$$

$$+ \begin{pmatrix} 0 \\ \widehat{\beta}_{(n)}\delta\alpha_{(n)} + \widehat{\alpha}_{(n)}\delta\beta_{(n)} + \widehat{\beta}_{(n)}\delta\beta_{(n)} \end{pmatrix} \simeq \begin{pmatrix} 0 \\ \tau_0^{n+3} \end{pmatrix}.$$

Since all three terms are bounded by $(\tau_0^3, \tau_0^{n+3})^{\mathsf{T}}$, so is their sum. This proves our assertion Eq. (15). □

Having showed its faithfulness, we are now confident in using the estimator Eq. (8) to study the full $\Omega_{(n)}$ of CDD_n. In particular, we have $\alpha_{(n)} \approx \widehat{\alpha}_{(n)} = 4^n\alpha$ and $\beta_{(n)} \approx \widehat{\beta}_{(n)}$. Similar as the PDD case, we can derive a strict bound and an approximate bound. For the strict bound, we have

$$\begin{aligned} \widehat{\beta}_{(n)} &\leq 2^{n(n+1)}\|\mathrm{ad}_{B_0}^n(B_1)\| \\ &\quad + 2^{n^2}\|\mathrm{ad}_{B_0}^{n-1}([B_0, B_2] - \mathrm{i}\{B_1, B_3\})\| \\ &\leq 2^{n(n+1)}\alpha^{n-1}\beta(2^n\alpha + \alpha + \beta). \end{aligned} \tag{19}$$

9

References

[1] Kaveh Khodjasteh and DA Lidar. Fault-tolerant quantum dynamical decoupling. *Physical review letters*, 95(18):180501, 2005.