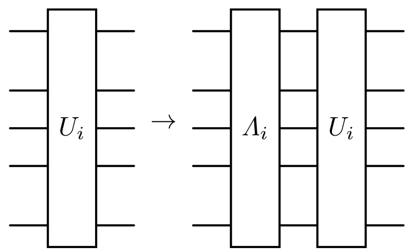
8.2.2 Probabilistic error cancellation

The Probabilistic error cancellation (PEC) method cancels the effects of the noise employing a map that acts as the inverse of the noise map under a suitable average.

Suppose the ideal circuit is performed by a unitary CPTP map $\mathcal{U}$ being the consecutive application of unitary gates: $\hat{U}_{\text{circuit}} = \hat{U}_d \dots \hat{U}_1$, where d is the depth of the circuit. One can represent the corresponding noisy circuit by substituting each unitary operation with its noisy counterpart, namely $\hat{U}_i \hat{\rho} \hat{U}_i^\dagger \to \hat{U}_i \Lambda_i[\hat{\rho}] \hat{U}_i^\dagger$, where Λ_i is a CPTP noisy map and $\hat{\rho}$ is the N-qubit state. If we focus on a single gate $\hat{U}_i$, the two corresponding circuits are represented as



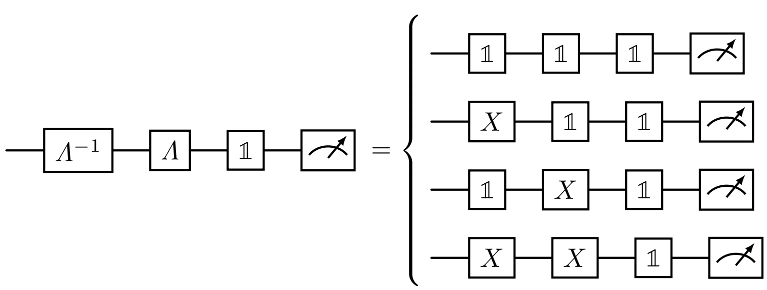
(8.98)

Now, the point is if we can invert the CPTP map Λ_i via the application of its inverse Λ_i^{-1}. In general, this is not possible. Indeed, typically, Λ_i^{-1} is not a CPTP map and thus such an inverse operation of the noise cannot be implemented. Nevertheless, such an operation can be implemented on average.

Consider the toy model of a single qubit, where the unitary noiseless operation is the identity: $\hat{U} = \hat{\mathbb{1}}$, and the noise channel is the bit-flip with a probability p. Thus, the corresponding map is

$$\Lambda(\hat{\rho}) = (1-p)\hat{\mathbb{1}}\hat{\rho}\hat{\mathbb{1}} + p\hat{\sigma}_x\hat{\rho}\hat{\sigma}_x. \tag{8.99}$$

This map corresponds to the unravelling with two components: with a probability p one applies an extra gate X, and with probability $(1-p)$ one does nothing, i.e. applies the gate $\mathbb{1}$. Notably, both these gates have an inverse. Indeed, $\hat{\mathbb{1}}^{-1} = \hat{\mathbb{1}}$ and $\hat{\sigma}_x^{-1} = \hat{\sigma}_x$. Then, we construct the inverse noise map Λ^{-1} as having two components: with a probability q we apply an X gate, and with a probability $(1-q)$ we apply an $\mathbb{1}$ gate. The corresponding total circuit can be then decomposed in the four components:



a), with prob $= (1-p)(1-q)$,

b), with prob $= (1-p)q$,

c), with prob $= p(1-q)$,

d), with prob $= pq$.

(8.100)

Now, we want to fix q such that, under the ensemble average, the circuit b) occurs with a probability being the opposite value of that of circuit c) occuring, and that the sum of the probabilities of having the circuit a) and d) gives 1. This implies the following system of equations

$$P_a P_d = (1-p)(1-q) + pq = 1, \quad \text{and} \quad P_b P_c = (1-p)q + p(1-q) = 0. \tag{8.101}$$

The solution is given by

$$q = \frac{-p}{1-2p}, \tag{8.102}$$

which is a quasi-probability, since it can take negative values, and it is shown in the left panel of Fig. 8.3. Now, the inverse noise map Λ^{-1} is given by

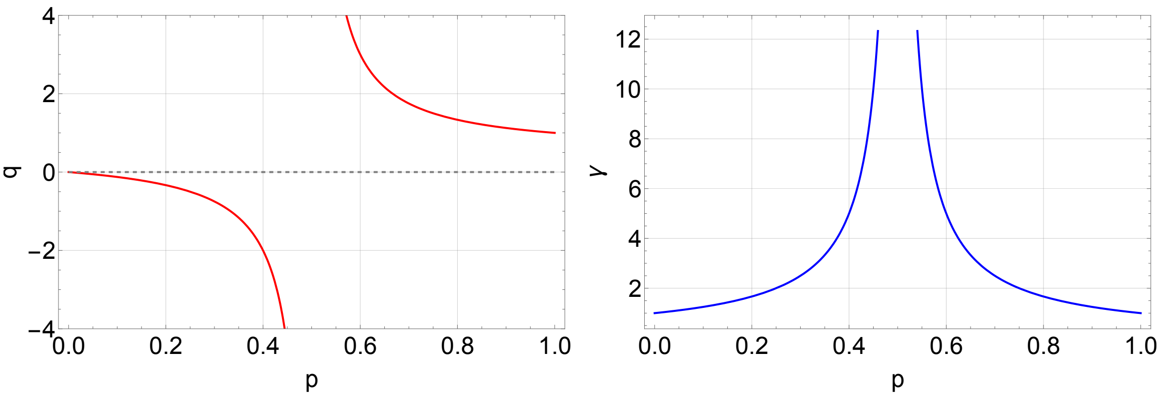


Fig. 8.3: Quasi-probability q (left panel) and renormalisation constant γ (right panel) as a function of the probability p of having an error.

$$\begin{aligned} \Lambda^{-1}(\hat{\rho}) &= (1-q)\hat{\mathbb{1}}\hat{\rho}\hat{\mathbb{1}} + q\hat{\sigma}_x\hat{\rho}\hat{\sigma}_x, \\ &= \text{sgn}(1-q)|1-q|\hat{\mathbb{1}}\hat{\rho}\hat{\mathbb{1}} + \text{sgn}(q)|q|\hat{\sigma}_x\hat{\rho}\hat{\sigma}_x, \\ &= \gamma\left[S_{\mathbb{1}}P_{\mathbb{1}}\hat{\mathbb{1}}\hat{\rho}\hat{\mathbb{1}} + S_X P_X \hat{\sigma}_x\hat{\rho}\hat{\sigma}_x\right], \end{aligned} \tag{8.103}$$

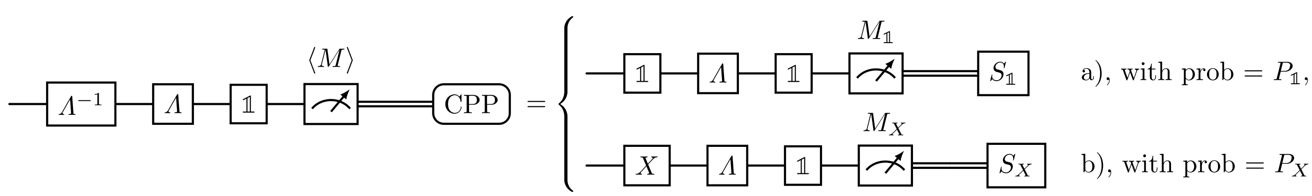
where in the first line we used $x = \text{sgn}(x)|x|$, with sgn indicating the sign function, and in the second line we introduced

$$\gamma = |1-q| + |q|, \tag{8.104}$$

which is represented in the right panel of Fig. 8.3. Finally, we defined

$$\begin{aligned} S_{\mathbb{1}} &= \text{sgn}(1-q), \quad P_{\mathbb{1}} = \frac{|1-q|}{\gamma}, \\ S_X &= \text{sgn}(q), \quad P_X = \frac{|q|}{\gamma}, \end{aligned} \tag{8.105}$$

where Thus, one has that $P_{\mathbb{1}}, P_X \geq 0$. Thus, independently from the unravelling of the noise map, i.e. without knowing if the bit-flip noise is applied or not, we apply the map Λ^{-1} as written in the last line of Eq. (8.103). This can be implemented with the following circuit:



(8.106)

A classical post-processing (CPP) is applied to multiplicate the outcome of the result by the proper sign factor. Eventually, the mitigated result is given by

$$\langle M \rangle = \gamma\left(S_{\mathbb{1}}P_{\mathbb{1}}M_{\mathbb{1}} + S_X P_X M_X\right). \tag{8.107}$$

This is an unbiased estimator. The cost of the mitigation procedure goes in the variance, which grows by a factor γ^2 compared to the unmitigated one.

Consider a more general case of the noise map Λ acting on a single qubit, which reads

$$\Lambda(\hat{\rho}) = \lambda_0\hat{\rho} + \lambda_1\hat{\sigma}_x\hat{\rho}\hat{\sigma}_x + \lambda_2\hat{\sigma}_y\hat{\rho}\hat{\sigma}_y + \lambda_3\hat{\sigma}_z\hat{\rho}\hat{\sigma}_z, \tag{8.108}$$

where

$$\lambda_\alpha \geq 0, \quad \text{and} \quad \sum_{\alpha=0}^{3}\lambda_\alpha = 1. \tag{8.109}$$

Such a map is a CPTP map. Similarly as done above, we construct the inverse map Λ^{-1} as

$$\Lambda^{-1}(\hat{\rho}) = q_0\hat{\rho} + q_1\hat{\sigma}_x\hat{\rho}\hat{\sigma}_x + q_2\hat{\sigma}_y\hat{\rho}\hat{\sigma}_y + q_3\hat{\sigma}_z\hat{\rho}\hat{\sigma}_z, \tag{8.110}$$

where we require that

$$\sum_{\alpha=0}^{3} q_\alpha = 1, \tag{8.111}$$

but we do not add any restriction on the sign of q_α. Then, in terms of unravellings, we have 4 possible evolutions provided by Λ and 4 by Λ^{-1} for a total of 16 possible mappings. Explicitly, they give

Λ	Λ^{-1}	$\hat{\rho} \to \hat{\rho}'$	probability $P_{\alpha\beta}$
$\hat{\mathbb{1}}$	$\hat{\mathbb{1}}$	$\hat{\rho}$	$\lambda_0 q_0$
$\hat{\mathbb{1}}$	$\hat{\sigma}_x$	$\hat{\sigma}_x\hat{\rho}\hat{\sigma}_x$	$\lambda_0 q_1$
$\hat{\mathbb{1}}$	$\hat{\sigma}_y$	$\hat{\sigma}_y\hat{\rho}\hat{\sigma}_y$	$\lambda_0 q_2$
$\hat{\mathbb{1}}$	$\hat{\sigma}_z$	$\hat{\sigma}_z\hat{\rho}\hat{\sigma}_z$	$\lambda_0 q_3$
$\hat{\sigma}_x$	$\hat{\mathbb{1}}$	$\hat{\sigma}_x\hat{\rho}\hat{\sigma}_x$	$\lambda_1 q_0$
$\hat{\sigma}_x$	$\hat{\sigma}_x$	$\hat{\sigma}_x^2\hat{\rho}\hat{\sigma}_x^2$	$\lambda_1 q_1$
$\hat{\sigma}_x$	$\hat{\sigma}_y$	$\hat{\sigma}_x\hat{\sigma}_y\hat{\rho}\hat{\sigma}_y\hat{\sigma}_x$	$\lambda_1 q_2$
$\hat{\sigma}_x$	$\hat{\sigma}_z$	$\hat{\sigma}_x\hat{\sigma}_z\hat{\rho}\hat{\sigma}_z\hat{\sigma}_x$	$\lambda_1 q_3$
$\hat{\sigma}_y$	$\hat{\mathbb{1}}$	$\hat{\sigma}_y\hat{\rho}\hat{\sigma}_y$	$\lambda_2 q_0$
$\hat{\sigma}_y$	$\hat{\sigma}_x$	$\hat{\sigma}_y\hat{\sigma}_x\hat{\rho}\hat{\sigma}_x\hat{\sigma}_y$	$\lambda_2 q_1$
$\hat{\sigma}_y$	$\hat{\sigma}_y$	$\hat{\sigma}_y^2\hat{\rho}\hat{\sigma}_y^2$	$\lambda_2 q_2$
$\hat{\sigma}_y$	$\hat{\sigma}_z$	$\hat{\sigma}_y\hat{\sigma}_z\hat{\rho}\hat{\sigma}_z\hat{\sigma}_y$	$\lambda_2 q_3$
$\hat{\sigma}_z$	$\hat{\mathbb{1}}$	$\hat{\sigma}_z\hat{\rho}\hat{\sigma}_z$	$\lambda_3 q_0$
$\hat{\sigma}_z$	$\hat{\sigma}_x$	$\hat{\sigma}_z\hat{\sigma}_x\hat{\rho}\hat{\sigma}_x\hat{\sigma}_z$	$\lambda_3 q_1$
$\hat{\sigma}_z$	$\hat{\sigma}_y$	$\hat{\sigma}_z\hat{\sigma}_y\hat{\rho}\hat{\sigma}_y\hat{\sigma}_z$	$\lambda_3 q_2$
$\hat{\sigma}_z$	$\hat{\sigma}_z$	$\hat{\sigma}_z^2\hat{\rho}\hat{\sigma}_z^2$	$\lambda_3 q_3$

(8.112)

However, we can exploit that $\hat{\sigma}_\alpha^2 = \hat{\mathbb{1}}$ and that $\hat{\sigma}_i\hat{\sigma}_j = i\epsilon_{ijk}\hat{\sigma}_k$. Thus, the above table becomes

$$
\begin{array}{c|c|c|c}
\Lambda & \Lambda^{-1} & \hat{\rho} \to \hat{\rho}' & \text{probability } P_{\alpha\beta} \\
\hline
\hat{\mathbb{1}} & \hat{\mathbb{1}} & \hat{\rho} & \lambda_0 q_0 \\
\hat{\mathbb{1}} & \hat{\sigma}_x & \hat{\sigma}_x\hat{\rho}\hat{\sigma}_x & \lambda_0 q_1 \\
\hat{\mathbb{1}} & \hat{\sigma}_y & \hat{\sigma}_y\hat{\rho}\hat{\sigma}_y & \lambda_0 q_2 \\
\hat{\mathbb{1}} & \hat{\sigma}_z & \hat{\sigma}_z\hat{\rho}\hat{\sigma}_z & \lambda_0 q_3 \\
\hat{\sigma}_x & \hat{\mathbb{1}} & \hat{\sigma}_x\hat{\rho}\hat{\sigma}_x & \lambda_1 q_0 \\
\hat{\sigma}_x & \hat{\sigma}_x & \hat{\rho} & \lambda_1 q_1 \\
\hat{\sigma}_x & \hat{\sigma}_y & \hat{\sigma}_z\hat{\rho}\hat{\sigma}_z & \lambda_1 q_2 \\
\hat{\sigma}_x & \hat{\sigma}_z & \hat{\sigma}_y\hat{\rho}\hat{\sigma}_y & \lambda_1 q_3 \\
\hat{\sigma}_y & \hat{\mathbb{1}} & \hat{\sigma}_y\hat{\rho}\hat{\sigma}_y & \lambda_2 q_0 \\
\hat{\sigma}_y & \hat{\sigma}_x & \hat{\sigma}_z\hat{\rho}\hat{\sigma}_z & \lambda_2 q_1 \\
\hat{\sigma}_y & \hat{\sigma}_y & \hat{\rho} & \lambda_2 q_2 \\
\hat{\sigma}_y & \hat{\sigma}_z & \hat{\sigma}_x\hat{\rho}\hat{\sigma}_x & \lambda_2 q_3 \\
\hat{\sigma}_z & \hat{\mathbb{1}} & \hat{\sigma}_z\hat{\rho}\hat{\sigma}_z & \lambda_3 q_0 \\
\hat{\sigma}_z & \hat{\sigma}_x & \hat{\sigma}_y\hat{\rho}\hat{\sigma}_y & \lambda_3 q_1 \\
\hat{\sigma}_z & \hat{\sigma}_y & \hat{\sigma}_x\hat{\rho}\hat{\sigma}_x & \lambda_3 q_2 \\
\hat{\sigma}_z & \hat{\sigma}_z & \hat{\rho} & \lambda_3 q_3
\end{array}
\tag{8.113}
$$

Finally, we impose that the sum of the probabilities of getting $\hat{\rho}' = \hat{\rho}$ should be 1, and those such $\hat{\rho}' \neq \hat{\rho}$ should be 0. Namely

$$
\begin{aligned}
P_{00} + P_{11} + P_{22} + P_{33} &= \lambda_0 q_0 + \lambda_1 q_1 + \lambda_2 q_2 + \lambda_3 q_3 = 1, \\
P_{01} + P_{10} + P_{23} + P_{32} &= \lambda_0 q_1 + \lambda_1 q_0 + \lambda_2 q_3 + \lambda_3 q_2 = 0, \\
P_{02} + P_{20} + P_{13} + P_{31} &= \lambda_0 q_2 + \lambda_2 q_0 + \lambda_1 q_3 + \lambda_3 q_1 = 0, \\
P_{03} + P_{30} + P_{12} + P_{21} &= \lambda_0 q_3 + \lambda_3 q_0 + \lambda_1 q_2 + \lambda_2 q_1 = 0.
\end{aligned}
\tag{8.114}
$$

The solution to this system of linear equations gives

$$
\begin{aligned}
q_0 &= \frac{1}{4}\left(1 + \frac{1}{1-2\lambda_1-2\lambda_2} + \frac{1}{1-2\lambda_1-2\lambda_3} + \frac{1}{1-2\lambda_2-2\lambda_3}\right), \\
q_1 &= \frac{1}{4}\left(1 - \frac{1}{1-2\lambda_1-2\lambda_2} - \frac{1}{1-2\lambda_1-2\lambda_3} + \frac{1}{1-2\lambda_2-2\lambda_3}\right), \\
q_2 &= \frac{1}{4}\left(1 - \frac{1}{1-2\lambda_1-2\lambda_2} + \frac{1}{1-2\lambda_1-2\lambda_3} - \frac{1}{1-2\lambda_2-2\lambda_3}\right), \\
q_3 &= \frac{1}{4}\left(1 + \frac{1}{1-2\lambda_1-2\lambda_2} - \frac{1}{1-2\lambda_1-2\lambda_3} - \frac{1}{1-2\lambda_2-2\lambda_3}\right).
\end{aligned}
\tag{8.115}
$$

The inverse map can be rewritten as

$$
\begin{aligned}
\Lambda^{-1}(\hat{\rho}) &= \sum_{\alpha=0}^{3} q_\alpha \hat{\sigma}_\alpha \hat{\rho} \hat{\sigma}_\alpha, \\
&= \sum_{\alpha=0}^{3} \mathrm{sgn}(q_\alpha)|q_\alpha| \hat{\sigma}_\alpha \hat{\rho} \hat{\sigma}_\alpha, \\
&= \gamma \sum_{\alpha=0}^{3} S_\alpha P_\alpha \hat{\sigma}_\alpha \hat{\rho} \hat{\sigma}_\alpha,
\end{aligned}
\tag{8.116}
$$

where

$$
\gamma = \sum_{\alpha=0}^{3} |q_\alpha|, \quad S_\alpha = \mathrm{sgn}(q_\alpha), \quad \text{and} \quad P_\alpha = \frac{|q_\alpha|}{\gamma}.
\tag{8.117}
$$

Then, the mitigated result is given by

$$\langle M \rangle = \gamma \sum_{\alpha=0}^{3} S_\alpha P_\alpha M_\alpha, \tag{8.118}$$

where M_α is the outcome obtained from the measurement at the end of the circuit at whose beginning we applied $\hat{\sigma}_\alpha$.

The application of PEC mitigation works if one has an almost perfect knowledge of the noise. However, for such a characterisation for N qubits, one needs to quantify $4^N - 1$ parameters, where 4 is the dimensions of the single-qubit algebra and 1 degree of freedom is fixed as it corresponds to the map given by $\hat{\mathbb{1}}^{\otimes N}$ whose associated probability is given by the unity minus the sum of all the other probabilities. To be quantitative, for 2 qubits one needs 15 parameters, for 10 qubits these become $\sim 10^6$, and for 50 qubits we have $\sim 10^{30}$ parameters. Therefore, it is an approach that requires too many classical processing to be used for a large number of qubits.