EPJ.org

○ **EPJ Quantum Technology**
a SpringerOpen Journal

Qubit noise deconvolution

Stefano Mangini[1,2*] ⓘ, Lorenzo Maccone[1,2,3] ⓘ and Chiara Macchiavello[1,2,3] ⓘ

*Correspondence:
stefano.mangini01@
universitadipavia.it
[1] Dipartimento di Fisica, Università di
Pavia, Via Bassi 6, I-27100, Pavia, Italy
[2] INFN Sezione di Pavia, Via Bassi 6,
I-27100, Pavia, Italy
Full list of author information is
available at the end of the article

Abstract

We present a noise deconvolution technique to remove a wide class of noises when performing arbitrary measurements on qubit systems. In particular, we derive the inverse map of the most common single qubit noisy channels, and exploit it at the data processing step to obtain noise-free estimates of observables evaluated on a qubit system subject to known noise. We illustrate a self-consistency check to ensure that the noise characterization is accurate providing simulation results for the deconvolution of a generic Pauli channel, as well as experimental evidence of the deconvolution of decoherence noise occurring on Rigetti quantum hardware.

1 Introduction

Quantum noise is currently the largest limiting factor in the adoption of quantum computation and quantum technology. Their theoretical performances are in fact hindered by the intrinsic fragility of quantum systems, and over the last years many proposal have been put forward to mitigate, ideally correct, the effect of noise and recover reliable results. On the computing side, as fault-tolerant quantum computers remain out of reach at the moment [1–4], various error mitigation techniques have been proposed to extend the capabilities of current small scale noisy quantum devices [5–7]. These ranges from correcting the readout noise via inversion of probability assignment matrix [8], extrapolating the noise in the device to the zero error case [9–11], using a probabilistic sampling on specific circuits to approximate the noise free computation [9, 10, 12], to also using machine learning approaches to learn how to recover ideal results [13].
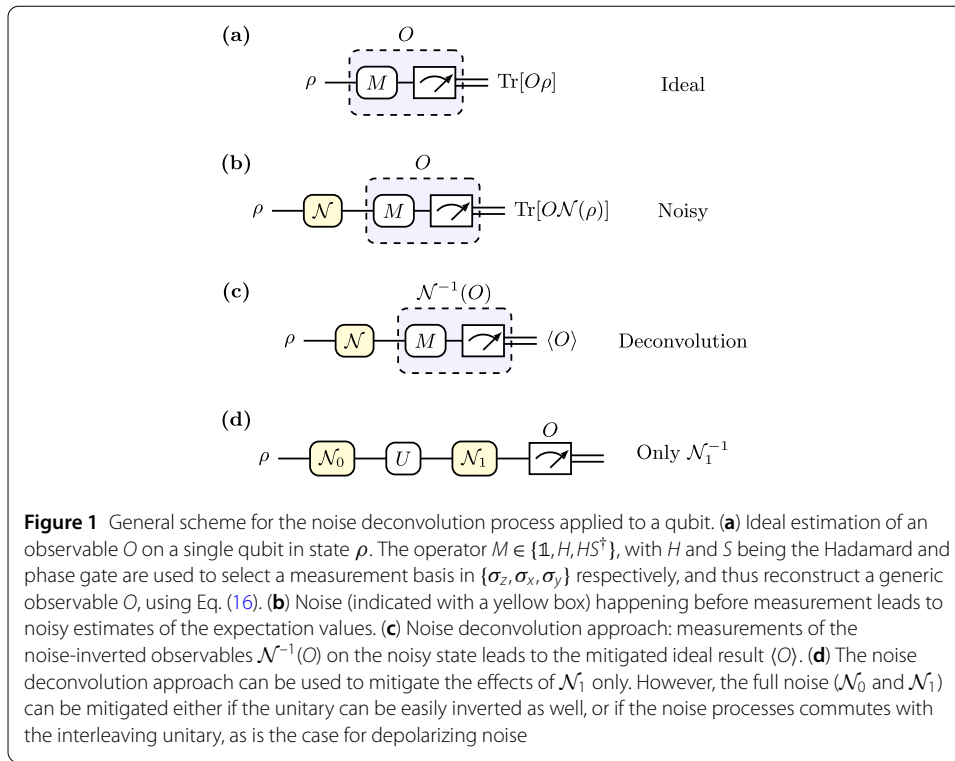
While these methods are concerned with mitigating noise occurring in a computation, here we instead focus on the more generic task of correcting the expectation value of arbitrary observables evaluated on a system which is subject to a known noise happening before the measurement stage. Such a scenario is relevant in quantum communication and quantum tomography tasks [14].

Noise in quantum systems is described by means of quantum channels [15]

$$\rho \longrightarrow \mathcal{E}(\rho) = \sum_k A_k \rho A_k^\dagger, \tag{1}$$

where A_k are operators acting on the system named Kraus operators. While the effect of unitary dynamics can be reversed using realizable operations, quantum channels cannot

⌖ Springer

Figure 1 General scheme for the noise deconvolution process applied to a qubit. (**a**) Ideal estimation of an observable O on a single qubit in state ρ. The operator $M \in \{\mathbb{1}, H, HS^\dagger\}$, with H and S being the Hadamard and phase gate are used to select a measurement basis in $\{\sigma_z, \sigma_x, \sigma_y\}$ respectively, and thus reconstruct a generic observable O, using Eq. (16). (**b**) Noise (indicated with a yellow box) happening before measurement leads to noisy estimates of the expectation values. (**c**) Noise deconvolution approach: measurements of the noise-inverted observables $\mathcal{N}^{-1}(O)$ on the noisy state leads to the mitigated ideal result $\langle O \rangle$. (**d**) The noise deconvolution approach can be used to mitigate the effects of $\mathcal{N}_1$ only. However, the full noise ($\mathcal{N}_0$ and $\mathcal{N}_1$) can be mitigated either if the unitary can be easily inverted as well, or if the noise processes commutes with the interleaving unitary, as is the case for depolarizing noise

be undone, and one can only hope to find operations which only approximately invert the noise process at hand. Examples of this approach leverages for example Petz recovery maps [16–18], or unitaries which, on average, are able to best reverse the noise based on given distance measures [19–21].

Here instead we show that noise can be eliminated by means of a *deconvolution* process, provided that the noise map describing the process is known and invertible. In fact, we drop the requirements of the inverse transformation being itself a quantum channel, since the transformation is not applied to the quantum system itself, but to the outcome statistics as a classical post-processing step. We derive the inversion maps of the most common single-qubit noisy channels (both unital and non-unital), and show how to use these to remove the effect of noise from the expectation values of general observables. In Fig. 1 we schematically summarize the noise deconvolution idea. The mitigation is effectively obtained by multiplying the noisy estimates by a factor depending on the noise $\langle O \rangle_{\text{mitig}} \sim c \langle O \rangle_{\text{noisy}}$, which comes at the cost of increasing the variance of the estimation, as $\text{Var}[\langle O \rangle_{\text{mitig}}] \sim c^2 \, \text{Var}[\langle O \rangle_{\text{noisy}}]$, so one needs to gather more statistic to reach a target precision. A related post processing technique specialized for quantum many-body systems and quantum field theory is put forward in ref. [22]. In addition, we provide both numerical simulations of the noise deconvolution process, as well as evidence of deconvolution of decoherence noise occurring on the superconducting quantum computer "Aspen-9" provided by Rigetti, accessed using the Quantum Cloud Services (QCS) [23]. We show how simple self-consistency checks can test whether the known noise map is accurate and how a feedback scheme can be used to adjust the noise parameters. Our contributions then include: (*i*) formalization and discussion of CPTP (namely, *completely positive trace preserving*) noise deconvolution of expectation values through (mathematical) inversion of the noise map; (*ii*) explicit derivation of the inverse map of the most common single

qubit noise channels; (*iii*) numerical and experimental application of the ideas introduced before.

Before continuing, we briefly describe the relation of the proposed noise deconvolution idea to probabilistic error mitigation (PEC) [10, 12], a quantum error mitigation technique aimed at correcting noisy operations during a quantum computation. Given a characterization of the noise, PEC works by using the inverse noise map of the operations to build an ensemble of suitably generated quantum circuits. These are sampled according to specific weights, and the results combined to build an approximation of the action of the noise-free quantum circuit. In particular, the mitigation procedure is *active*, in the sense that the experimenter need to generate new quantum circuits and run them against the quantum device. On the contrary, we are instead concerned with the correction of expectation values evaluated on a noisy state, with no computation or dynamics involved. In addition, within our framework, the mitigation is *passive*, in the sense that the mitigation happens classically as a post-processing step, and no action on the quantum system is necessary. Appropriately limiting PEC to the specific case of measurement error mitigation, and realizing that sampling on quantum circuits is no longer a necessary step, then one can recover the noise deconvolution procedure presented here, whose regime of application is not restricted to quantum computation, but applies to a general quantum mechanical measurement scenario. As such, some of the results presented here can be recovered also with the techniques proposed in [10, 12]. That said, the explicit calculations presented here for the general noise maps we analyze have not been presented elsewhere in full generality, e.g. see Table 1 below.

The paper is organized as follows. In Sect. 2 we recall some basic concepts about quantum channels and the Pauli transfer matrix formalism, and the idea of noise deconvolution in Sect. 3. In Sect. 4 we leverage the Pauli transfer matrix formalism to explicitly derive the inverse map of the most common single qubit noise channels, and use inside the noise deconvolution procedure to obtain noise-free estimates. In Table 1 we summarize all the maps taken in consideration as well as their inverse. In Sect. 5 we show by means of simulations that the noise deconvolution process can be used to cancel out the effect of a general Pauli channel, and also provide experimental evidence of the deconvolution of decoherence noise as performed on a real quantum device by Rigetti.

2 Methods

In this section we introduce the notation and the theoretical tools used to derive the main results of the work. We will denote with $\mathcal{H}$ the Hilbert space, and with $\mathcal{L}(\mathcal{H})$ the space of squared linear operators acting on $\mathcal{H}$. For those interested, a brief overview of quantum channels and Kraus decomposition can be found in Appendix A.

2.1 Quantum channels

In general quantum channels cannot be physically inverted, as there is no quantum evolution capable of reversing their actions. Formally stated, let $\mathcal{E}$ be a CPTP map, it is not possible to find another CPTP map $\mathcal{D} = \mathcal{E}^{-1}$, such that $(\mathcal{D} \circ \mathcal{E})(\rho) = \rho \ \forall \rho$. The only trivial case when this is possible, is for maps having only a single Kraus operator, in which case they reduce to standard unitary evolution $\mathcal{E}(\rho) = U\rho U^{\dagger}$, with the inverse given by $\mathcal{D}(\cdot) = U^{\dagger}(\cdot)U$.

The CPTP conditions impose hard constraints to the operatorial form that physically realizable evolutions must match, namely the Kraus representation. However, the requirement for admitting a more general operator-sum representation are looser. In fact, any Hermiticity preserving map, i.e. a map such that $\Phi(\rho)^{\dagger} = \Phi(\rho)$ for $\rho = \rho^{\dagger}$, admits an operator-sum representation as [24, 25]

$$\Phi(\rho) = \sum_k \lambda_k A_k \rho A_k^{\dagger} \tag{2}$$

with $\lambda_k \in \{+1, -1\}$. Clearly if all the coefficients are $\lambda_k = 1$ $\forall k$, then the map Φ is also completely positive, since it is in standard Kraus form (1). Moreover, another useful characterization is given by

Corollary 1 (Corollary II.2 of [25]) *Let $\mathcal{M}_N$ be the space of complex $N \times N$ matrices. Suppose $\Phi : \mathcal{M}_N \to \mathcal{M}_N$ is a completely positive map having the form*

$$\Phi(\rho) = \sum_k \beta_k A_k \rho A_k^{\dagger}, \tag{3}$$

where $\{A_k\}_k$ is linearly independent in $\mathcal{M}_N$, and $\beta_k \in \mathbb{R}$ $\forall k$. Then $\beta_k \geq 0$ $\forall k$.

Conversely, if a map has form (3) with linearly independent operators $\{A_k\}_k$ but has *some* of the coefficients $\beta_j < 0$, then the map is not completely positive. This result is steadily applied to maps acting on qubit systems where $\mathcal{M}_N = \mathbb{C}^{2\times2}$. In fact, Pauli matrices σ_x, σ_y and σ_z together with the identity $\sigma_0 = \mathbb{1}_2$, form a linearly independent set in the space of 2×2 complex matrices, and then any map of the form

$$\mathcal{E}(\rho) = \beta_0 \sigma_0 \rho \sigma_0 + \beta_1 \sigma_x \rho \sigma_x + \beta_2 \sigma_y \rho \sigma_y + \beta_3 \sigma_z \rho \sigma_z \tag{4}$$

having some negative coefficients is not a CP map, thus it is not a physically realizable channel. In the following we will derive many inverse maps having this form, for which this result holds. Of course, we already know that a quantum channel cannot be inverted (apart from the trivial unitary case), so that if an inversion map is found, then it is certainly not CP. Nonetheless, this result is still of interest because it gives a nice and clear condition that can be used to quickly assess the nature of the maps under investigation. In addition, as shown in ref. [26], if a CPTP map is invertible, then its inverse is Hermitian preserving (HP), and so can be expressed in operator-form (2).

2.2 Qubit systems and Pauli transfer matrix formalism

We focus on quantum systems made of qubits, and briefly review some useful results on qubit channels.

The identity and the Pauli matrices $\{\mathbb{1}, \sigma_x, \sigma_y, \sigma_z\}$ form a basis on $\mathcal{L}(\mathcal{H}) = \mathbb{C}^{2\times2}$, and so any operator $O \in \mathcal{L}(\mathcal{H})$ can be expressed in this basis as

$$O = \sum_{i=0}^{3} o_i \sigma_i, \quad o_i = \mathrm{Tr}[\sigma_i O] \tag{5}$$

$$= o_0 \mathbb{1} + \boldsymbol{o} \cdot \boldsymbol{\sigma}, \tag{6}$$

where we have introduced the vector of Pauli matrices $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$, and the vector of coefficients $\boldsymbol{o} = (o_1, o_2, o_3) \in \mathbb{C}^3$. Similarly, density operators are expressed in this basis in terms of their Bloch vector as $\rho = (\mathbb{1} + \boldsymbol{r} \cdot \boldsymbol{\sigma})/2$, with $\boldsymbol{r} = (r_x, r_y, r_z) \in \mathbb{R}^3$, and $|\boldsymbol{r}| \leq 1$, where equality holds only for pure states $\rho = |\psi\rangle\langle\psi|$ [15].

Since any operator O is completely specified by its components in the Pauli basis, we define its vector of coefficients as the column vector $|O\rangle\!\rangle := (o_0, o_1, o_2, o_3)^\mathsf{T}$. We refer to [27, 28] for a detailed discussion on the correspondence between operators and vectors.

In addition, every linear map $\Phi : \mathcal{L}(\mathbb{C}^2) \to \mathcal{L}(\mathbb{C}^2)$ can be represented in this basis as a 4×4 matrix Γ [25, 29–31], whose action is given by

$$
\begin{aligned}
\Phi(O) \longrightarrow \Gamma|O\rangle\!\rangle &= \begin{bmatrix} \gamma_0 & \boldsymbol{\gamma} \\ \boldsymbol{t} & T \end{bmatrix} \begin{bmatrix} o_0 \\ \boldsymbol{o} \end{bmatrix} \\
&= \begin{bmatrix} \gamma_0 o_0 + \boldsymbol{\gamma} \cdot \boldsymbol{o} \\ o_0 \boldsymbol{t} + T\boldsymbol{o} \end{bmatrix} \quad \text{or}
\end{aligned}
\tag{7}
$$

$$
\Phi(O) = (\gamma_0 o_0 + \boldsymbol{\gamma} \cdot \boldsymbol{o})\mathbb{1} + (o_0 \boldsymbol{t} + T\boldsymbol{o}) \cdot \boldsymbol{\sigma},
$$

where $\boldsymbol{\gamma}$ and $\boldsymbol{t}$ are row and column vectors respectively, and T is a 3×3 matrix. The matrix Γ associated to the map Φ is called Pauli Transfer Matrix (PTM), and its elements are given by

$$
\Gamma_{ij} = \frac{1}{2} \operatorname{Tr}\big[\sigma_i \Phi(\sigma_j)\big], \quad i, j \in \{0, 1, 2, 3\}, \sigma_0 = \mathbb{1}.
\tag{8}
$$

If we restrict to trace-preserving maps, then $\boldsymbol{\gamma} = \boldsymbol{0}$ and $\gamma_0 = 1$, so the Γ matrix reduces to the simpler form

$$
\Gamma = \begin{bmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{t} & T \end{bmatrix}.
\tag{9}
$$

Furthermore, if the map is also unital (i.e. $\Phi(\mathbb{1}) = \mathbb{1}$), then also $\boldsymbol{t} = \boldsymbol{0}$. As an example, the quantum bit-flip channel described by the map

$$
\mathcal{N}_x(\rho) = (1 - p)\rho + p\sigma_x \rho \sigma_x,
\tag{10}
$$

has a corresponding PTM representation as

$$
\mathcal{N}_x \longrightarrow \Gamma_x = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 - 2p & 0 \\ 0 & 0 & 0 & 1 - 2p \end{bmatrix}.
\tag{11}
$$

2.3 Quantum tomographic reconstruction
Quantum tomography [32–35] is a method to estimate the ensemble average of any arbitrary operator by using measurement outcomes of a quorum of observables. The goal of a tomographic reconstruction of an observable is to identify a set of observables $\{Q_\lambda\}$, called *quorum* [36], such that the mean value $\langle O \rangle = \operatorname{Tr}[O\rho]$ of any observable $O \in \mathcal{L}(\mathcal{H})$,

for all states ρ, can be reconstructed by using measurements outcomes of the quorum observables. A tomographic reconstruction formula for an operator O is obtained by using a spectral decomposition of the identity in the operator Hilbert space [36–39]

$$O = \int_\Lambda d\lambda \, \text{Tr}\big[C_\lambda^\dagger O\big]C_\lambda, \tag{12}$$

where λ is a parameter living in either a continuous or discrete manifold Λ, and operators C_λ depend on the quorum observables. The term $\mathbb{E}[O](Q_\lambda) := \text{Tr}[C_\lambda^\dagger O]C_\lambda$ is called *quantum estimator* of the operator O, and given a quantum state ρ, the expectation value $\langle O \rangle$ on such state amounts to

$$\langle O \rangle = \text{Tr}[O\rho] = \int_\Lambda d\lambda \, \text{Tr}\big[OC_\lambda^\dagger\big]\text{Tr}[C_\lambda\rho] \tag{13}$$

$$= \int_\Lambda d\lambda \, \text{Tr}\big[\mathbb{E}[O](Q_\lambda)\rho\big] \tag{14}$$

$$= \int_\Lambda d\lambda \, \big\langle \mathbb{E}[O](Q_\lambda)\big\rangle. \tag{15}$$

For qubit systems, the most common choice (but non unique, e.g. [35]) for the quorum are the Pauli matrices $\{Q_\lambda\}_\lambda = \{\sigma_x, \sigma_y, \sigma_z\}$, and the tomographic reconstruction formula results in the standard expansion in the Pauli basis, albeit with a slightly different notation (see Appendix B for the explicit derivation):

$$\langle O \rangle = \sum_{\alpha=x,y,z} \frac{1}{3}\big\langle\mathbb{E}[O](\sigma_\alpha)\big\rangle,$$

$$\mathbb{E}[O](\sigma_\alpha) = \left(\frac{3\,\text{Tr}[O\sigma_\alpha]}{2}\sigma_\alpha + \frac{\text{Tr}[O]}{2}\mathbb{1}\right). \tag{16}$$

Note that the quantum tomographic reconstruction can be straightforwardly applied to multipartite quantum systems by simply using as a quorum the tensor product of single-system quorums [36].

3 Noise deconvolution

The tomographic reconstruction formula can be used whenever one has access to the quantum state ρ and measurements of the quorum observables. In practical scenarios however, estimations are performed in the presence of noise and one generally deals with noisy quantum states $\rho \to \tilde{\rho} = \mathcal{N}(\rho)$ which then leads to noisy estimates $\langle O \rangle_{\tilde{\rho}} = \text{Tr}[O\mathcal{N}(\rho)]$. The idea of noise deconvolution is to correct the errors by considering a new quorum of observables taking into account the noise, and then use a noise inverted quantum estimator to recover the ideal estimates, namely the ones that we would obtain in the absence of noise.

Suppose the noise map $\mathcal{N}$ acting on the quantum system can be formally inverted, that is there exist a linear (not CP) map $\mathcal{N}^{-1}$ such that $(\mathcal{N}^{-1} \circ \mathcal{N})(\rho) = \rho$ for all ρ. Then, we say that the noise can be *deconvoluted* in the following sense: instead of measuring the original observable O, we can evaluate the expectation value of the noise-inverted operator

$\hat{\mathcal{N}}^{-1}(O)$, thus obtaining as a result the desired noise-free ideal result $\langle O \rangle$, that is

$$
\begin{aligned}
\langle \hat{\mathcal{N}}^{-1}(O) \rangle_{\tilde{\rho}} &= \mathrm{Tr}\big[\hat{\mathcal{N}}^{-1}(O)\mathcal{N}(\rho) \big] \\
&= \mathrm{Tr}\big[O\mathcal{N}^{-1}\big(\mathcal{N}(\rho) \big) \big] \\
&= \mathrm{Tr}[O\rho] = \langle O \rangle,
\end{aligned}
\tag{17}
$$

where $\hat{\mathcal{N}}^{-1}(\cdot)$ denotes the adjoint of the inverse map $\mathcal{N}^{-1}(\cdot)$, and in the second line we made explicit use of the definition of the adjoint map.

The conditions for *deconvolving* the effect of a noise channel $\mathcal{N}$ are [36, 37]:

- the inverted noise map exists, that is there is a $\mathcal{N}^{-1}$ such that $(\mathcal{N}^{-1} \circ \mathcal{N})(O) = O$ $\forall O \in \mathcal{L}(\mathcal{H})$.
- the quantum estimator $\mathbb{E}[O](Q_\lambda)$ is in the domain of $\mathcal{N}^{-1}$.
- the map $\mathcal{N}^{-1}(\mathbb{E}[O](Q_\lambda))$ is a function of Q_λ.

If these hold, then one can substitute the quantum estimator in Eq. (13), with the deconvolved quantum estimator $\hat{\mathcal{N}}^{-1}(\mathbb{E}[O](Q_\lambda))$, yielding

$$
\begin{aligned}
&\int_\Lambda d\lambda \, \mathrm{Tr}\big[\hat{\mathcal{N}}^{-1}\big(\mathbb{E}[O](Q_\lambda) \big) \mathcal{N}(\rho) \big] \\
&= \int_\Lambda d\lambda \, \mathrm{Tr}\big[\mathbb{E}[O](Q_\lambda)\mathcal{N}^{-1}\big(\mathcal{N}(\rho) \big) \big] \\
&= \int_\Lambda d\lambda \, \mathrm{Tr}\big[\mathbb{E}[O](Q_\lambda)\rho \big] \\
&= \mathrm{Tr}[O\rho] = \langle O \rangle.
\end{aligned}
\tag{18}
$$

This procedure yields the ideal expectation value of any observable O on the state ρ, even if having access only to a noisy version of it and provided that the noise map is known (and invertible). Note that this definition is similar to that recently reported in ref [6], regarding invertible noise channels with non-CPTP inverse. Specializing it for qubits, using Eq. (18) in (16), leads to (see Appendix C for further details)

$$
\langle O \rangle = \frac{1}{2}\mathrm{Tr}[O] + \frac{1}{2}\sum_{\alpha=x,y,z}\mathrm{Tr}[O\sigma_\alpha]\langle \hat{\mathcal{N}}^{-1}(\sigma_\alpha) \rangle_{\tilde{\rho}}.
\tag{19}
$$

Similarly to standard tomographic reconstruction, noise deconvolution can be applied also to multi qubits systems, in which case the mitigated tomographic estimates are obtained considering the tensor product of the deconvolved quantum estimator of each subsystem. Clearly, this holds only for single-qubit noisy channels acting independently on each qubit. In addition, generally non-invertible maps could still be deconvoluted if one restricts the attention only to a subset of states of interest upon which the given map is invertible [40, 41].

As shown later, the correction of the expectation value of a Pauli matrix is obtained by multiplying the noisy estimate—the one the experimenter has access to—by a constant depending on the noise, i.e. $\langle \sigma_\alpha \rangle_{\mathrm{mitig}} = c\langle \sigma_\alpha \rangle_{\mathrm{noisy}}$. This clearly increases the variance of the estimation, since $\mathrm{Var}[\langle \sigma_\alpha \rangle_{\mathrm{mitig}}] = c^2\,\mathrm{Var}[\langle \sigma_\alpha \rangle_{\mathrm{noisy}}] \sim c^2/M$, where M is the number of measurements performed on the system, and thus the experimenter need to increase the outcome statistics proportionally to c^2 to reach a desired target precision.

We now proceed discussing how the deconvolution behaves in the presence of multiple noise channels. Consider two noise processes $\mathcal{N}_0$ and $\mathcal{N}_1$ separated by a unitary gate $\mathcal{U}(\cdot) = U \cdot U^\dagger$, as shown in Fig. 1(d). The action of the circuit is

$$(\mathcal{N}_1 \circ \mathcal{U} \circ \mathcal{N}_0)(\rho) = \mathcal{N}_1\big(U\mathcal{N}_0(\rho)U^\dagger\big)$$
$$= \mathcal{N}_1(\tilde{\rho}_U)$$

with $\tilde{\rho}_U = U\mathcal{N}_0(\rho)U^\dagger$. Using (18), it is possible to deconvolve the outermost noise $\mathcal{N}_1$ with

$$\mathrm{Tr}\big[\hat{\mathcal{N}}_1^{-1}(O)\mathcal{N}_1(\tilde{\rho}_U)\big], \tag{20}$$

but not $\mathcal{N}_0$, since the unitary U is in the way. Actually, one could decide to deconvolve the unitary as well, using the trivial inverse $\mathcal{U}_1^{-1}(\cdot) = U_1^\dagger \cdot U_1$, and thus making it possible to deconvolve also the first noise channel $\mathcal{N}_0$, as

$$\mathrm{Tr}\big[\hat{\mathcal{N}}_1^{-1}\big((U\hat{\mathcal{N}}_0^{-1}(O)U^\dagger)\mathcal{N}_1(\tilde{\rho}_U)\big]$$
$$= \mathrm{Tr}\big[(U\hat{\mathcal{N}}_0^{-1}(O)U^\dagger)U\mathcal{N}_0(\rho)U^\dagger)\big]$$
$$= \mathrm{Tr}\big[\hat{\mathcal{N}}_0^{-1}(O)\mathcal{N}_0(\rho)\big]$$
$$= \mathrm{Tr}[O\rho].$$

However, this procedure cannot be employed to invert the noise that happens before a generic unitary U, since it essentially offloads the computation from the quantum computer to the classical one, by simulating the inverse evolution of the quantum system.

A more interesting case is obtained when the error map happens to commute with all the remaining operation in the computation, as is the case for the depolarizing noise, described by the map

$$\mathcal{N}_{\mathrm{dep}}(\rho) = \frac{p\mathbb{1}}{2} + (1-p)\rho, \tag{21}$$

for which it is easy to see that $(\mathcal{N}_{\mathrm{dep}} \circ \mathcal{U})(\rho) = (\mathcal{U} \circ \mathcal{N}_{\mathrm{dep}})(\rho)\ \forall \mathcal{U}(\cdot) = U \cdot U^\dagger$. Suppose one is performing a quantum computation given by a sequence of operations U_i, each one followed by depolarizing noise

$$\rho = \left(\prod_{i=1}^{d} \mathcal{N}_{\mathrm{dep}}^{(i)} \circ \mathcal{U}_i\right)(\rho_0) \tag{22}$$

$$= \left(\prod_{i=1}^{d} \mathcal{N}_{\mathrm{dep}}^{(i)} \circ \prod_{i=1}^{d} \mathcal{U}_i\right)(\rho_0) \tag{23}$$

$$= \mathcal{N}_{\mathrm{dep}}^{\mathrm{tot}}(\rho_U), \tag{24}$$

with $\mathcal{N}_{\mathrm{dep}}^{\mathrm{tot}} = \prod \mathcal{N}_{\mathrm{dep}}^{(i)}$ the composition of all the depolarizing channels, and $\rho_U = \prod \mathcal{U}_i(\rho_0)$ the state obtained by the ideal noise-free computation. Most importantly, one can check that the composition of multiple depolarizing channels is still a depolarizing channel with probability parameter $1 - p_{\mathrm{tot}} = \prod(1 - p_i)$, where p_i is the probability associated with each

depolarizing noise. In such case it is possible to deconvolve all noise at once, using the deconvolution formula for the depolarizing noise with the total noise parameter p_{tot} (see Eq. (39)). Similarly, this also holds for computations involving multi qubits subject to *global* depolarizing errors. The authors in ref. [42] leverage this property to perform a simple yet effective error mitigation technique for quantum computers, based on the assumption that noise in quantum circuits is well described by global depolarizing error channels. While exact depolarizing errors (either local or global) are hardly found in realistic quantum circuits where errors are both due to *coherent* (i.e. unitary) and *incoherent* noise (i.e. interaction), Pauli twirling and randomized compiling techniques [43–46] can be used to approximately tailor noise to stochastic Pauli channels, preferably depolarizing noise, and then use the procedure above to mitigate it [47].

4 Inversion of common noise maps

We now proceed by explicitly evaluating the inverse maps of some of the most common noisy channels, leveraging the Pauli Transfer Matrix formalism introduced in Sect. 2. The general method for finding the inverse map goes as follows: we first evaluate the matrix representation (7) of the channel, we then invert this matrix, and from this recover the operator sum representation of the inverse channel whenever this exists. We start from simpler cases to build some intuition on the construction of the inverse maps, and then proceed towards more complicated cases. In Table 1 we summarize the results obtained in this section, comprising all noise channels considered in this analysis together with their inverse maps.

4.1 Bit-flip, phase-flip and bit-phase-flip

The bit-flip, phase-flip and bit-phase-flip channels are described by the Kraus operators, $A_0 = \sqrt{p}\mathbb{1}$ and $A_{1,\alpha} = \sqrt{1-p}\sigma_\alpha$, with $\sigma_\alpha \in \{\sigma_x, \sigma_z, \sigma_y\}$ respectively. For simplicity, in the

Table 1 Table summarizing the results of the present analysis, consisting of some of the most common single-qubit noisy channels $\mathcal{N}$, along with their inverse noise maps $\mathcal{N}^{-1}$, defined as the map such that $(\mathcal{N}^{-1} \circ \mathcal{N})(\rho) = \rho \; \forall \rho$. Clearly, all noise channels are CPTP maps, while the inverse channels are not, yet they admit an operator-sum representation. All the noise maps except for amplitude damping and 2-Kraus channel have trivial adjoint channels, so one must pay attention in using the adjoint channel inside the deconvolution formula (19)

	Noise channel $\mathcal{N}(\rho)$	Inverse map $\mathcal{N}^{-1}(O)$
Bit-Flip	$(1-p)\rho + p\sigma_x\rho\sigma_x$	$\frac{1-p}{1-2p}O - \frac{p}{1-2p}\sigma_x O\sigma_x$
Phase-Flip (or *dephasing*)	$(1-p)\rho + p\sigma_z\rho\sigma_z$	$\frac{1-p}{1-2p}O - \frac{p}{1-2p}\sigma_z O\sigma_z$
Bit-Phase-Flip	$(1-p)\rho + p\sigma_y\rho\sigma_y$	$\frac{1-p}{1-2p}O - \frac{p}{1-2p}\sigma_y O\sigma_y$
Depolarizing	$(1-p)\rho + p\frac{\mathbb{1}}{2}$	$\frac{1}{1-p}(O - \frac{p}{2}\,\text{Tr}[O]\mathbb{1})$
General Pauli Channel	$p_0\rho + p_x\sigma_x\rho\sigma_x + p_y\sigma_y\rho\sigma_y + p_z\sigma_z\rho\sigma_z$	$\beta_0 O + \beta_1\sigma_x O\sigma_x + \beta_2\sigma_y O\sigma_y + \beta_3\sigma_z O\sigma_z$ (see Eq. (41) for the coefficients)
Amplitude Damping	$V_0\rho V_0 + V_1\rho V_1^\dagger$	$K_0 O K_0 - K_1 O K_1^\dagger$
	$V_0 = \|0\rangle\langle 0\| + \sqrt{1-\gamma}\|1\rangle\langle 1\|$	$K_0 = \|0\rangle\langle 0\| + \sqrt{\frac{1}{1-\gamma}}\|1\rangle\langle 1\|$
	$V_1 = \sqrt{\gamma}\|0\rangle\langle 1\|$	$K_1 = \sqrt{\frac{\gamma}{1-\gamma}}\|0\rangle\langle 1\|$
2-Kraus Channel	$A_0\rho A_0 + A_1\rho A_1^\dagger$	$B_1 O B_1^\dagger - B_2 O B_2^\dagger$
	$A_0 = \cos\alpha\|0\rangle\langle 0\| + \cos\beta\|1\rangle\langle 1\|$	$B_0 = \frac{\sqrt{2}\cos\beta}{\sqrt{\cos 2\alpha + \cos 2\beta}}\|0\rangle\langle 0\| + \frac{\sqrt{2}\cos\alpha}{\sqrt{\cos 2\alpha + \cos 2\beta}}\|1\rangle\langle 1\|$
	$A_1 = \sin\beta\|0\rangle\langle 1\| + \sin\alpha\|1\rangle\langle 0\|$	$B_1 = \frac{\sqrt{2}\sin\beta}{\sqrt{\cos 2\alpha + \cos 2\beta}}\|0\rangle\langle 1\| + \frac{\sqrt{2}\sin\alpha}{\sqrt{\cos 2\alpha + \cos 2\beta}}\|1\rangle\langle 0\|$

following we focus only on the bit-flip channel (generated by σ_x), but the results hold equivalently also for the other two channels. The bit-flip channel acts as:

$$\mathcal{N}_x(\rho) = (1 - p)\rho + p\sigma_x\rho\sigma_x \tag{25}$$

and its PTM is given by

$$\Gamma_x = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 - 2p & 0 \\ 0 & 0 & 0 & 1 - 2p \end{bmatrix}. \tag{26}$$

In order to find an operator sum expression for the inverse map $\mathcal{N}_x^{-1}$, we consider the inverse matrix Γ_x^{-1}

$$\Gamma_x^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{(1-2p)} & 0 \\ 0 & 0 & 0 & \frac{1}{(1-2p)} \end{bmatrix}. \tag{27}$$

It's clear that Γ_x can be inverted provided that $p \neq 1/2$, since in that case $\det \Gamma_x = 0$. This is not a problem for realistic case scenarios, where the probability of errors are small, so that one can safely assume $0 < p < 1/2$. We now proceed using a similar procedure found in [25].

Note that Γ_x^{-1} is diagonal in the Pauli basis, thus has eigenvectors $\{|\mathbb{1}\rangle\!\rangle, |\sigma_x\rangle\!\rangle |\sigma_y\rangle\!\rangle |\sigma_z\rangle\!\rangle\}$ with eigenvalues $\boldsymbol{\lambda} = \{1, 1, (1 - 2p)^{-1}, (1 - 2p)^{-1}\}$ respectively. Now, consider the generic map

$$\mathcal{E}(O) = \sum_{j=0}^{3} \beta_j \sigma_j O \sigma_j. \tag{28}$$

Also this map has eigenvectors $\{\mathbb{1}, \boldsymbol{\sigma}\}$, but with eigenvalues $\boldsymbol{\beta} = \{\beta_0, \beta_1, \beta_2, \beta_3\}$. Since two maps are equal if they have the same action on a basis, if we can find a way to match the two sets of eigenvalues $\boldsymbol{\lambda}$ and $\boldsymbol{\beta}$, we would then recover the operator-sum representation for Γ_x^{-1}.

By evaluating the PTM $\Gamma_{\mathcal{E}}$ of $\mathcal{E}$, we can relate the coefficients in the operator-sum representation (28), with those appearing in the expression for Γ_x^{-1} (see Appendix D for a derivation). In particular, we want these to hold:

$$\left(\Gamma_x^{-1}\right)_{11} = \beta_0 + \beta_1 - \beta_2 - \beta_3, \tag{29}$$

$$\left(\Gamma_x^{-1}\right)_{22} = \beta_0 - \beta_1 + \beta_2 - \beta_3, \tag{30}$$

$$\left(\Gamma_x^{-1}\right)_{33} = \beta_0 - \beta_1 - \beta_2 + \beta_3 \tag{31}$$

plus the trace-preserving condition $1 = \beta_0 + \beta_1 + \beta_2 + \beta_3$, that the inverse map must satisfy because the direct map is trace-preserving. This condition is inherently satisfied by Γ_x^{-1},

since its first row has the form $(1, 0, 0, 0)$. This system has solutions $\beta_0 = (1 - p)/(1 - 2p)$, $\beta_1 = -p/(1 - 2p)$, and $\beta_2 = \beta_3 = 0$, and substituting them back into Eq. (28), we obtain the operator-sum representation of the inverse bit-flip map

$$\mathcal{N}_x^{-1}(O) = \frac{1-p}{1-2p}O - \frac{p}{1-2p}\sigma_x O \sigma_x. \tag{32}$$

By virtue of Corollary II, and noticing that the coefficients appearing in the expression above have always opposite signs, we are sure that this map is not CP, as expected, yet it possesses an operator sum representation. Note how similar the direct and inverse map are, a feature which we will encounter in all the cases discussed here.

The same procedure can be applied to phase-flip (or *dephasing*, generated by σ_z), and bit-phase-flip (generated by σ_y) channels, yielding inverse maps

$$\mathcal{N}_z^{-1}(O) = \frac{1-p}{1-2p}O - \frac{p}{1-2p}\sigma_z O \sigma_z, \tag{33}$$

$$\mathcal{N}_y^{-1}(O) = \frac{1-p}{1-2p}O - \frac{p}{1-2p}\sigma_y O \sigma_y. \tag{34}$$

We can plug these inversion maps in the deconvolution formula (19) to obtain noise-free expectation values. In particular, assume we are measuring a Pauli matrix $O = \sigma_\alpha$, and that the system is subject to one of the noise processes $\rho \rightarrow \rho_\beta = \mathcal{N}_\beta(\rho)$ with $\beta = \{x, y, z\}$. Then the ideal expectation values $\langle \sigma_\alpha \rangle_\rho = \text{Tr}[\sigma_\alpha \rho]$ can be expressed in compact form as (see Appendix D for the explicit derivation)

$$\langle \sigma_\alpha \rangle = \delta_{\alpha\beta} \langle \sigma_\alpha \rangle_{\rho_\beta} + (1 - \delta_{\alpha\beta}) \frac{1}{1-2p} \langle \sigma_\alpha \rangle_{\rho_\beta}. \tag{35}$$

It is then clear that if the noise happens along the measurement direction ($\alpha = \beta$), then the noise does not affect the measurement statistics, as the ideal and noisy value coincide. While for orthogonal directions ($\alpha \neq \beta$), these are equally contracted by a factor $1 - 2p$, thus recovering the usual pictorial representation of the contracting Bloch sphere on the plane orthogonal to the noise [15].

4.2 Depolarizing noise

The depolarizing noise, introduced above,

$$\mathcal{N}_{\text{dep}}(\rho) = (1 - p)\rho + \frac{p\mathbb{1}}{2}, \tag{36}$$

leaves the state untouched with probability $1 - p$, and sends it to the completely mixed state $\mathbb{1}/2$ with probability p. The channel can be expressed in Kraus form in multiple ways, one of them being [15]

$$\mathcal{N}_{\text{depol}}(\rho) = \left(1 - \frac{3p}{4}\right)\rho + \frac{p}{4}(\sigma_x \rho \sigma_x + \sigma_y \rho \sigma_y + \sigma_z \rho \sigma_z), \tag{37}$$

with Kraus operators $\{A_0 = \sqrt{1 - 3p/4}\mathbb{1}, A_1 = \sqrt{p}\sigma_x/2, A_2 = \sqrt{p}\sigma_y/2, A_3 = \sqrt{p}\sigma_z/2\}$. Following the same procedure used for the bit-flip channel, one obtains the inverse linear map

(see Appendix D)

$$\mathcal{N}_{\text{depol}}^{-1}(O) = \frac{1}{1-p}\left(O - \frac{p}{2}\text{Tr}[O]\mathbb{1}\right). \tag{38}$$

While this is already a known result in the literature [10, 36, 39, 48, 49], it is presented without an explicit constructive derivation, as given here.

Using this formula in the deconvolution tomographic reconstruction (19), we find

$$\langle O \rangle = \frac{1}{2}\text{Tr}[O] + \sum_\alpha \frac{\text{Tr}[O\sigma_\alpha]}{1-p}\langle\sigma_\alpha\rangle_{\mathcal{N}_{\text{dep}}(\rho)}, \tag{39}$$

where it is clear that to counterbalance the effect of the depolarizing channel, whose effect on the Bloch sphere is to contract it uniformly, one needs perform an expansion of the same amount, obtained dividing by $1-p$.

While our analysis is focused only on single qubit systems, it is worth noticing that a similar approach can be used to correct correlated and asymmetric depolarizing channels acting on multi-qubits systems [50], as recently shown in [27].

4.3 General Pauli channel

A more general and interesting case is the general Pauli channel, where noise intensities are different along the three Pauli axes

$$\mathcal{N}_{\boldsymbol{p}}(\rho) = p_0 O + p_x \sigma_x \rho \sigma_x + p_y \sigma_y \rho \sigma_y + p_z \sigma_z \rho \sigma_z. \tag{40}$$

The channel is parametrized by the probabilities $\boldsymbol{p} = (p_0, p_x, p_y, p_z)$, with the trace-preserving condition implying $p_0 = 1 - p_x - p_y - p_z$. Upon choosing appropriate values for $\boldsymbol{p}$, this channel reduces to all noise maps treated before. Though of considerable more general structure, the inverse map of this channel is derived using the same machinery developed above, and eventually one obtains

$$\mathcal{N}_{\boldsymbol{p}}^{-1}(O) = \beta_0 O + \beta_1 \sigma_x O \sigma_x + \beta_2 \sigma_y O \sigma_y + \beta_3 \sigma_z O \sigma_z \quad \text{with}$$

$$\beta_0 = \frac{1}{4}\left(1 + \frac{1}{1-2(p_x+p_y)} + \frac{1}{1-2(p_x+p_z)} + \frac{1}{1-2(p_y+p_z)}\right),$$

$$\beta_1 = \frac{1}{4}\left(1 - \frac{1}{1-2(p_x+p_y)} - \frac{1}{1-2(p_x+p_z)} + \frac{1}{1-2(p_y+p_z)}\right),$$

$$\beta_2 = \frac{1}{4}\left(1 - \frac{1}{1-2(p_x+p_y)} + \frac{1}{1-2(p_x+p_z)} - \frac{1}{1-2(p_y+p_z)}\right),$$

$$\beta_3 = \frac{1}{4}\left(1 + \frac{1}{1+2(p_x+p_y)} - \frac{1}{1-2(p_x+p_z)} - \frac{1}{1-2(p_y+p_z)}\right). \tag{41}$$

One can check that varying $\boldsymbol{p}$ it is possible to recover the inverse maps of all the cases treated before. For example, for $\boldsymbol{p} = (1-p, p, 0, 0)$ corresponding to the bit-flip channel, one gets $\beta_0 = (1-p)/(1-2p)$ and $\beta_1 = -p/(1-2p)$, as in Eq. (32).

The noise deconvolution applied to measurements of Pauli matrices $O \in \{\sigma_x, \sigma_y, \sigma_z\}$, leads to the following relations

$$
\begin{aligned}
\langle \sigma_x \rangle &= \frac{1}{1 - 2(p_y + p_z)} \langle \sigma_x \rangle_{\mathcal{N}_{\mathbf{p}}(\rho)}, \\
\langle \sigma_y \rangle &= \frac{1}{1 - 2(p_x + p_z)} \langle \sigma_y \rangle_{\mathcal{N}_{\mathbf{p}}(\rho)}, \\
\langle \sigma_z \rangle &= \frac{1}{1 - 2(p_x + p_y)} \langle \sigma_z \rangle_{\mathcal{N}_{\mathbf{p}}(\rho)},
\end{aligned}
\tag{42}
$$

which can be used together with (19) to reconstruct the expectation value of a general observable O. Again, we see that the noise disturbs the estimation along orthogonal directions, as in all previous cases. Note that the explicit inversion of the general Pauli channel was also recently reported in ref. [5].

4.4 Amplitude damping

The amplitude damping (AD) channel describes the energy loss of a quantum system, for example obtained through relaxation from the excited to the ground state. Its Kraus representation is

$$
\begin{aligned}
\mathcal{N}_{\mathrm{AD}}(\rho) &= V_0 \rho V_0^\dagger + V_1 \rho V_1^\dagger, \\
V_0 &= \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix}, \qquad V_1 = \begin{bmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{bmatrix},
\end{aligned}
\tag{43}
$$

where $\gamma \in [0, 1]$ is a parameter that encodes the strength of the energy loss process, which for real systems is often expressed in terms of characteristic decay times, as discussed in Sect. 5.

While still being trace preserving (TP), amplitude damping channel is not unital, since $\mathcal{N}_{\mathrm{AD}}(\mathbb{1}) = \mathbb{1} + \gamma Z$. This in turn implies that the Pauli Transfer Matrix Γ_{AD} is not diagonal, but has an addition nonzero element in the last row of first column. This changes the derivation of the inverse map with respect to the previous cases, but it can still be carried out without major changes (see Appendix E). The inverse linear map in operator sum representation is then found to be

$$
\begin{aligned}
\mathcal{N}_{\mathrm{AD}}^{-1}(\rho) &= K_0 O K_0^\dagger - K_1 O K_1^\dagger, \\
K_0 &= \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\sqrt{1-\gamma}} \end{bmatrix}, \qquad K_1 = \begin{bmatrix} 0 & \sqrt{\frac{\gamma}{1-\gamma}} \\ 0 & 0 \end{bmatrix}.
\end{aligned}
\tag{44}
$$

Up until now, all noisy channels (and their inverse maps) had trivial *adjoint* map, since all Kraus operators were Hermitian. However this is not the case for amplitude damping, since both $V_1 \neq V_1^\dagger$ and $K_1 \neq K_1^\dagger$. Thus, one must be careful in applying the adjoint inverse $\hat{\mathcal{N}}^{-1}$ in Eq. (19), and not just $\mathcal{N}^{-1}$ of (44) (see Appendix E for an extended discussion).

Deconvolution of amplitude damping for measurements of the Pauli matrices leads to

$$
\begin{aligned}
\langle \sigma_x \rangle &= \frac{1}{\sqrt{1-\gamma}} \langle \sigma_x \rangle_{\mathcal{N}_{\mathrm{AD}}(\rho)}, \\
\langle \sigma_y \rangle &= \frac{1}{\sqrt{1-\gamma}} \langle \sigma_y \rangle_{\mathcal{N}_{\mathrm{AD}}(\rho)}, \\
\langle \sigma_z \rangle &= \frac{1}{1-\gamma} \big(\langle \sigma_z \rangle_{\mathcal{N}_{\mathrm{AD}}(\rho)} - \gamma \big).
\end{aligned}
\tag{45}
$$

Similarly, one can also obtain the inverse map of the *generalized amplitude damping* (GAD) channel, used to model the interaction of a qubit with an environment at a finite temperature [15, 51]. Such channel is parametrised by two parameters γ and p, and it is defined as

$$
\begin{aligned}
\mathcal{N}_{\mathrm{GAD}}(\rho) &= A_0 \rho A_0^\dagger + A_1 \rho A_1^\dagger + A_2 \rho A_2^\dagger + A_3 \rho A_3^\dagger, \\
A_0 &= \sqrt{p} \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix}, \qquad A_1 = \sqrt{p} \begin{bmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{bmatrix}, \\
A_2 &= \sqrt{1-p} \begin{bmatrix} \sqrt{1-\gamma} & 0 \\ 0 & 1 \end{bmatrix}, \qquad A_3 = \sqrt{1-p} \begin{bmatrix} 0 & 0 \\ \sqrt{\gamma} & 0 \end{bmatrix}.
\end{aligned}
\tag{46}
$$

One can check that the following map is the inverse of the GAD channel

$$
\begin{aligned}
\mathcal{N}_{\mathrm{GAD}}^{-1}(\rho) &= B_0 \rho B_0^\dagger - B_1 \rho B_1^\dagger + B_2 \rho B_2^\dagger - B_3 \rho B_3^\dagger, \\
B_0 &= \sqrt{p} \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{\frac{1}{1-\gamma}} \end{bmatrix}, \qquad B_1 = \sqrt{p} \begin{bmatrix} 0 & \sqrt{\frac{\gamma}{1-\gamma}} \\ 0 & 0 \end{bmatrix}, \\
B_2 &= \sqrt{1-p} \begin{bmatrix} \sqrt{\frac{1}{1-\gamma}} & 0 \\ 0 & 1 \end{bmatrix}, \qquad B_3 = \sqrt{1-p} \begin{bmatrix} 0 & 0 \\ \sqrt{\frac{\gamma}{1-\gamma}} & 0 \end{bmatrix},
\end{aligned}
\tag{47}
$$

with corresponding noise deconolved Pauli expectation values given by

$$
\begin{aligned}
\langle \sigma_x \rangle &= \frac{1}{\sqrt{1-\gamma}} \langle \sigma_x \rangle_{\mathcal{N}_{\mathrm{GAD}}(\rho)}, \\
\langle \sigma_y \rangle &= \frac{1}{\sqrt{1-\gamma}} \langle \sigma_y \rangle_{\mathcal{N}_{\mathrm{GAD}}(\rho)}, \\
\langle \sigma_z \rangle &= \frac{1}{1-\gamma} \big(\langle \sigma_z \rangle_{\mathcal{N}_{\mathrm{GAD}}(\rho)} - \gamma(2p-1) \big).
\end{aligned}
\tag{48}
$$

4.5 Two-Kraus channels

We now move our attention to the set of channels generated by two parametrized Kraus operators

$$
\mathcal{N}_{\mathrm{two}}(\rho) = \sum_{i=1,2} A_i \rho A_i^\dagger,
\tag{49}
$$

with $A_1 = \cos\alpha |0\rangle\langle 0| + \cos\beta |1\rangle\langle 1|$, and $A_2 = \sin\beta |0\rangle\langle 1| + \sin\alpha |1\rangle\langle 0|$. This channel reduces to bit-flip for $\alpha = \beta$, and to amplitude damping for $\alpha = 0$.

Following a procedure similar to the amplitude damping case, the inverse map of the two-Kraus channels is found to be

$$\mathcal{N}_{\text{two}}(O)^{-1} = B_1 O B_1^{\dagger} - B_2 O B_2^{\dagger} \quad \text{with}$$

$$B_1 = \begin{bmatrix} \frac{\sqrt{2}\cos\beta}{\sqrt{\cos 2\alpha + \cos 2\beta}} & 0 \\ 0 & \frac{\sqrt{2}\cos\alpha}{\sqrt{\cos 2\alpha + \cos 2\beta}} \end{bmatrix}$$

$$B_2 = \begin{bmatrix} 0 & \frac{\sqrt{2}\sin\beta}{\sqrt{\cos 2\alpha + \cos 2\beta}} \\ \frac{\sqrt{2}\sin\alpha}{\sqrt{\cos 2\alpha + \cos 2\beta}} & 0 \end{bmatrix}$$

(50)

Similarly to amplitude damping, one of the generators (B_2) is not Hermitian, thus we must employ the adjoint inverse map when evaluating the deconvolved expectation values. By straightforward calculations the following holds:

$$\langle \sigma_x \rangle = \frac{1}{\cos(\alpha - \beta)} \langle \sigma_x \rangle_{\mathcal{N}_{\text{two}}(\rho)},$$

$$\langle \sigma_y \rangle = \frac{1}{\cos(\alpha + \beta)} \langle \sigma_y \rangle_{\mathcal{N}_{\text{two}}(\rho)},$$

(51)

$$\langle \sigma_z \rangle = h_{\alpha\beta} \left(\cos^2\beta + \sin^2\alpha - 1 + \langle \sigma_z \rangle_{\mathcal{N}_{\text{two}}(\rho)} \right),$$

with $h_{\alpha\beta} = 2/(\cos(2\alpha) + \cos(2\beta))$.

Note that upon varying the parameters α and β, the formulas above correctly reduce to the other limiting channels. For example, setting $\alpha = 0$ leads to amplitude damping channel (45), with $\cos(\beta) := \sqrt{1 - \gamma}$.

5 Experimental deconvolution

In this section we provide some concrete applications of the noise deconvolution procedures for qubit tomography outlined above. In particular, we show both numerically and by experimentation on quantum hardware by Rigetti how to address a *dechoerence* noise model, and we also provide numerical evidence for the deconvolution of the general Pauli channel (40). All simulations are performed using PyQuil and the real quantum device usded is "Aspen-9", accessed via Rigetti's Quantum Cloud Services (QCS) [23, 52].

5.1 Decoherence noise model

The concurrent action of a dephasing channel followed by amplitude damping is referred to as *decoherence* noise, which is an effective way to describe the noisy evolution a qubit undergoes due to uncontrolled interaction with its external environment. Using the definitions (25) and (43), one obtains

$$\begin{aligned}
\mathcal{N}_{\text{dec}}(\rho) &= \left(\mathcal{N}_{\text{AD}}(\gamma) \circ \mathcal{N}_z(p) \right) \left(\begin{bmatrix} a & b \\ c & 1-a \end{bmatrix} \right) \\
&= \begin{bmatrix} 1 - (1-a)(1-\gamma) & (1-2p)\sqrt{1-\gamma}\,b \\ (1-2p)\sqrt{1-\gamma}\,c & (1-\gamma)(1-a) \end{bmatrix} \\
&= \begin{bmatrix} 1 - (1-a)e^{-t/T_1} & e^{-t/T_2}b \\ e^{-t/T_2}c & e^{-t/T_1}(1-a) \end{bmatrix},
\end{aligned}$$

(52)

where we have introduced the relaxation times T_1 and T_2 characterizing the "quality" of the physical qubits. These are related to the noise parameters γ and p through the following relations

$$\gamma = 1 - e^{-t/T_1},$$
$$p = \frac{1}{2}\left(1 - e^{-(t/T_2 - t/2T_1)}\right),$$

(53)

where t is a time parameter indicating the duration of the noise process.

Since the correction terms in the deconvolution formulas for dephasing (45) and amplitude damping (35) are multiplicative, for a decoherence channel these combine as
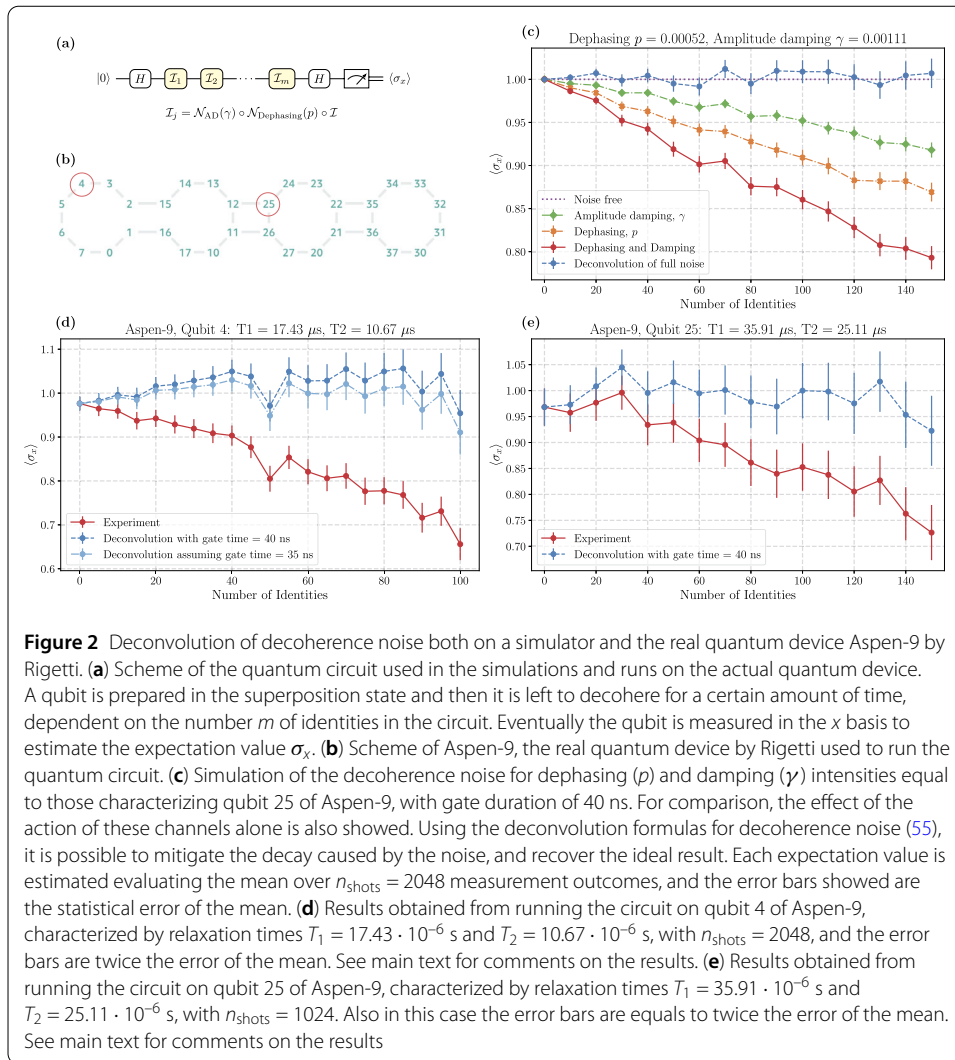
$$\langle \sigma_x \rangle = \frac{1}{(1 - 2p)} \frac{1}{\sqrt{1 - \gamma}} \langle \sigma_x \rangle_{\mathcal{N}_{\text{dec}}(\rho)},$$
$$\langle \sigma_y \rangle = \frac{1}{(1 - 2p)} \frac{1}{\sqrt{1 - \gamma}} \langle \sigma_y \rangle_{\mathcal{N}_{\text{dec}}(\rho)},$$
$$\langle \sigma_z \rangle = \frac{1}{1 - \gamma}\left(\langle \sigma_z \rangle_{\mathcal{N}_{\text{dec}}(\rho)} - \gamma\right).$$

(54)

Additionally, if the quantum system under investigation is subject to repeated applications of a decoherence noise channel, i.e. $\mathcal{N}_{\text{dec}}^{\circ m}(\rho) = \mathcal{N}^{(1)}{}_{\text{dec}} \circ \mathcal{N}_{\text{dec}}^{(2)} \cdots \circ \mathcal{N}_{\text{dec}}^{(m)}(\rho)$, then the ideal expectation values are obtained through the following equations

$$\langle \sigma_x \rangle = \frac{1}{((1 - 2p)\sqrt{1 - \gamma})^m} \langle \sigma_x \rangle_{\mathcal{N}_{\text{dec}}(\rho)},$$
$$\langle \sigma_y \rangle = \frac{1}{((1 - 2p)\sqrt{1 - \gamma})^m} \langle \sigma_y \rangle_{\mathcal{N}_{\text{dec}}(\rho)},$$
$$\langle \sigma_z \rangle = \frac{1}{(1 - \gamma)^m}\left(\langle \sigma_z \rangle_{\mathcal{N}_{\text{dec}}(\rho)} - 1 + (1 - \gamma)^m\right).$$

(55)

In Fig. 2 we show the application of these formulas to deconvolve the decoherence noise occurring on a qubit. The specific quantum circuit used for the experiments is showed in Fig. 2(a): first the system is prepared in the superposition state $|+\rangle = H|0\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, then we let qubit decohere for a certain amount of time dictated by the number m of (noisy) identities each of which takes a time t, and at last we measure the expectation value of the operator σ_x. Clearly, in a noise-free scenario, the result would always be $\langle \sigma_x \rangle = 1$, independent of the depth m. Figure 2(c) shows a simulation of these circuits with stochastic measurement outcomes for different values of m, and for a given choice of noise parameters p and γ. For comparison, the individual effect of dephasing and amplitude damping channels alone are also showed. Thanks to Eq. (55) we can invert the effect of the decoherence noise, and so retrieve the ideal noise-free results.

We also tested this procedure on quantum hardware provided by Rigetti, in particular on the device "Aspen-9", whose topology is reported in Fig. 2(b). The device comes with the calibration data reporting the T_1 and T_2 parameters for any qubit, as well as the time duration of a single gate. Identities in the circuits are used to introduce time delays, and thus let the qubit decohere for longer intervals of time, depending on the depth m. Differently from the previous simulations where only the identities are supposed to introduce (decoherence) noise, in the real case scenario noise happens along the whole computation, including state preparation, application of all gates in the circuit (both Hadamards and

Figure 2 Deconvolution of decoherence noise both on a simulator and the real quantum device Aspen-9 by Rigetti. (**a**) Scheme of the quantum circuit used in the simulations and runs on the actual quantum device. A qubit is prepared in the superposition state and then it is left to decohere for a certain amount of time, dependent on the number m of identities in the circuit. Eventually the qubit is measured in the x basis to estimate the expectation value σ_x. (**b**) Scheme of Aspen-9, the real quantum device by Rigetti used to run the quantum circuit. (**c**) Simulation of the decoherence noise for dephasing (p) and damping (γ) intensities equal to those characterizing qubit 25 of Aspen-9, with gate duration of 40 ns. For comparison, the effect of the action of these channels alone is also showed. Using the deconvolution formulas for decoherence noise (55), it is possible to mitigate the decay caused by the noise, and recover the ideal result. Each expectation value is estimated evaluating the mean over n_{shots} = 2048 measurement outcomes, and the error bars showed are the statistical error of the mean. (**d**) Results obtained from running the circuit on qubit 4 of Aspen-9, characterized by relaxation times $T_1 = 17.43 \cdot 10^{-6}$ s and $T_2 = 10.67 \cdot 10^{-6}$ s, with n_{shots} = 2048, and the error bars are twice the error of the mean. See main text for comments on the results. (**e**) Results obtained from running the circuit on qubit 25 of Aspen-9, characterized by relaxation times $T_1 = 35.91 \cdot 10^{-6}$ s and $T_2 = 25.11 \cdot 10^{-6}$ s, with n_{shots} = 1024. Also in this case the error bars are equals to twice the error of the mean. See main text for comments on the results

Identities), and finally measurement errors. Of these, the most detrimental are undoubtedly readout errors, and we addressed them by using the standard mitigation technique of calibrating the device and inverting the assignment probability matrix to recover readout mitigated results. Calibration data reports that the time it takes to execute a single qubit identity gate is t = 40 ns, and together with T_1 and T_2, these are used to calculate the parameters p and γ of the decoherence noise, using relations (53). These are in turn used inside the deconvolution formulas to recover the noise-free results. Figures 2(d) and 2(e) show the results of the execution of circuit Fig. 2(a) on qubits 4 and 25, respectively.

The noise mitigation procedure on qubit 4 shown in panel 2(d), yields slightly unphysical results, in the sense that the mitigated value exceeds one, which is of course not possible. A naive solution to this problem could be to impose that the mitigation results are in the physical range [−1, +1], so that if the result exceed the limits, it should be substituted with the appropriate physical bound. Though, assuming a gate time duration of t = 35 ns instead of standard 40 ns, yields results which are more in agreement with the expected theoretical behaviour for decoherence noise, as the deconvoluted results are compatible with one, as expected. This hints that either the quality of the qubit is better then reported in the available calibration data (either due to shorter gate times t, or larger T_1 and T_2),

or that the decoherence model alone poorly describes the noise happening on idle qubit 4 left interacting with the external environment. However, the good accordance between the deconvoluted results with $t = 35$ ns and the experiments suggests the first hypothesis to hold.

Such conclusion is also corroborated by the experimental results obtained with qubit 25. In fact, using the deconvolution formulas with reported T_1, T_2 and standard gate time ($t = 40$ ns), we are able to mitigate the effect of noise with good accuracy, as showed in Fig. 2(e), hinting that indeed the decay law of the qubit is well described through a decoherence noise model. Also, note that the simulation in Fig. 2(c) is tuned with the same noise parameters p and γ characterizing qubit 25. Apart from fluctuations due to, e.g imperfect readout, stochastic measurement outcomes, and noisy Hadamards, there is good agreement between the simulated (red curve in panel (c)) and experimental result (red curve in panel (e)). We do not report analogues experiments using other qubits in the device that produced obviously biased data.

5.2 Arbitrary Pauli channel

We implemented a simulation of the noise deconvolution of the general Pauli channel (40), using the quantum virtual machine (QVM) simulator provided with PyQuil [52]. The sim-
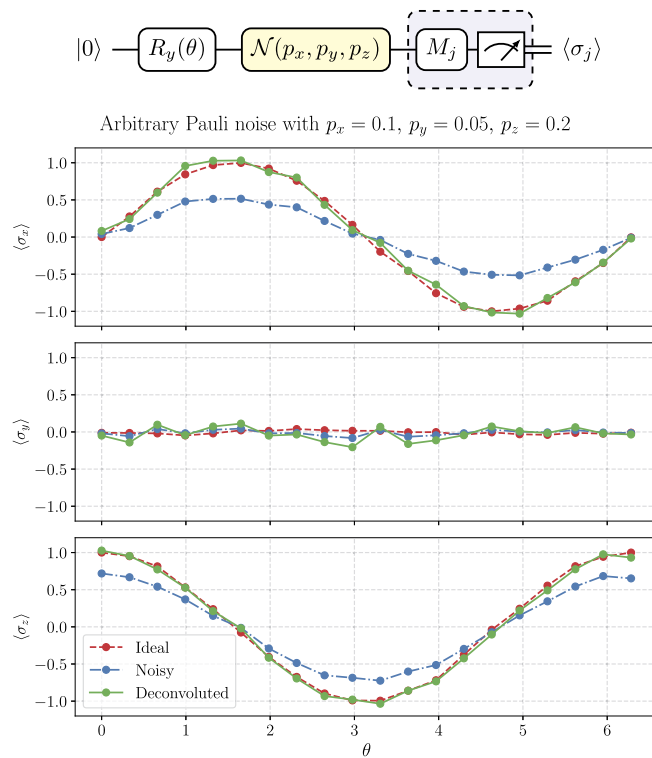
Figure 3 Simulation of the deconvolution process for the general Pauli channel $\mathcal{N}_{\boldsymbol{p}}$ (40). The noise parameters along the three Pauli axes are set to $p_x = 0.1$, $p_y = 0.05$, $p_z = 0.2$. The results are obtained simulating the circuit portrayed on top of the image for $n_{\text{shots}} = 1024$ shots and for multiple values of the angle θ. Then, the deconvolution formulas (42) are used to retrieve the ideal noise-free result. It is clear that the deconvolution effectively mitigates the Pauli noise yielding a final result which is much closer to the ideal noise-free one, up to differences due to stochastic measurement outcomes. In particular, the estimation of $\boldsymbol{\sigma}_y$ is dominated by the statistical error, which is amplified by the correction factor $1/(1 - 2(p_x + p_z)) = 2.5$

ulated circuit is showed on top of Fig. 3. A qubit starting in the ground state is rotated in the Bloch sphere around the y axis via $R_y(\theta) = e^{-i\theta\sigma_y/2}$, and then it is subject to the general Pauli noise (yellow box), simulated applying a Pauli transformation chosen randomly with probabilities p_x, p_y and p_z. At last, we estimate the expectation value of the three Pauli matrices by appending the appropriate change of basis gate, i.e $M_j \in \{\mathbb{1}, H, HS^\dagger\}$ for $\{\sigma_z, \sigma_x, \sigma_y\}$ respectively. The noise parameters (p_x, p_y, p_z) are used within the deconvolution formulas (42) to recover the mitigated results (green curve), which are, as expected, in perfect agreement with the ideal noise-free ones, obtained from executing the quantum circuit without the noisy channel (red curve).

6 Conclusions

In conclusion we have shown how mathematically invertible noise maps can always be removed from the final measurement stage, so that one can obtain unbiased expectation values of general observables provided that the noise process in known. We illustrated the method on most known qubit noise maps, and systematically derived their inverse maps (see Table 1). We simulated the noise deconvolution procedure for the case of a general Pauli channel (Fig. 3) and illustrated our method on noise on actual quantum hardware (Fig. 2).

Appendix A: Kraus decomposition

A quantum physical evolution is represented by *(i) linear, (ii) completely-positive* and *(iii) trace-preserving* (CPTP) maps taking quantum density operators to quantum density operators. A map satisfying these three properties is called a *quantum channel*, and can be interpreted as a quantum evolution obtained through the interaction of the system with an external environment. A map is a quantum channel if and only if it admits a Kraus (or *operator-sum*) representation as

$$\rho \longrightarrow \mathcal{E}(\rho) = \sum_k A_k \rho A_k^\dagger, \tag{A.1}$$

with the trace preserving condition requiring

$$\mathrm{Tr}\,\mathcal{E}(\rho) = \mathrm{Tr}\,\rho \quad \implies \quad \sum_k A_k^\dagger A_k = \mathbb{1}. \tag{A.2}$$

The operators $\{A_k\}_k$ are called the Kraus operators of the channel, which are however non-unique [15]. Such channels are often referred to as *stochastic* channels [29, 30], and if they also preserve the identity ($\mathcal{E}(\mathbb{1}) = \mathbb{1}$), they are called *unital* (or *bistochastic*). Unitality corresponds to the requirement that also $\sum_k A_k A_k^\dagger = \mathbb{1}$, from which it is clear that a sufficient condition for a CPTP map to be unital is for its Kraus operators to be self-adjoint $A_k = A_k^\dagger \,\forall k$.

Appendix B: Tomographic reconstruction formula for qubits

In this appendix we show how the tomographic reconstruction formula for systems made of qubits $\mathcal{H} = \mathbb{C}^2$ can be recovered starting from the standard basis expansion in terms of the Pauli matrices [35]. The set of matrices $\{\mathbb{1}, \sigma_x, \sigma_y, \sigma_z\}$ form an orthonormal set, and

constitutes a basis for the space of 2×2 complex matrices $\mathcal{L}(\mathcal{H}) = \mathbb{C}^{2 \times 2}$. So, given an operator $O \in \mathcal{L}(\mathcal{H})$, the following holds:

$$
\begin{aligned}
O &= \frac{\mathbb{1}\operatorname{Tr}[O] + \sigma_x \operatorname{Tr}[O\sigma_x] + \sigma_y \operatorname{Tr}[O\sigma_y] + \sigma_z \operatorname{Tr}[O\sigma_z]}{2} \\
&= \frac{\operatorname{Tr}[O]}{2}\mathbb{1} + \sum_{\alpha=x,y,z} \frac{\operatorname{Tr}[O\sigma_\alpha]}{2}\sigma_\alpha \\
&= \sum_{\alpha=x,y,z} \frac{1}{3}\left(\frac{3\operatorname{Tr}[O\sigma_\alpha]}{2}\sigma_\alpha + \frac{\operatorname{Tr}[O]}{2}\mathbb{1}\right) \\
&= \sum_{\alpha=x,y,z} \frac{1}{3}\mathcal{E}[O](\sigma_\alpha),
\end{aligned}
\tag{B.1}
$$

where

$$
\mathcal{E}[O](\sigma_\alpha) = \left(\frac{3\operatorname{Tr}[O\sigma_\alpha]}{2}\sigma_\alpha + \frac{\operatorname{Tr}[O]}{2}\mathbb{1}\right)
\tag{B.2}
$$

is the desired *quantum estimator*, with $\{\sigma_x, \sigma_y, \sigma_z\}$ constituting the *quorum* of observables of the tomographic reconstruction. Equation (B.1) has the same form of the tomographic reconstruction formula in Eq. (12), with substitutions

$$
\int_\Lambda \longrightarrow \sum_\lambda,
$$
$$
d\lambda \longrightarrow 1/3,
$$
$$
\lambda \longrightarrow \{x, y, z\},
$$
$$
\{Q_\lambda\} \longrightarrow \{\sigma_x, \sigma_y, \sigma_z\},
$$

which account for the fact that we are dealing with a discrete, and not continuous, basis expansion.

Also, note that (B.1) is not the unique choice for the tomographic formula. In fact, one could use a continuous parametrization of the group $SU(2)$ given by operator $\mathcal{R}(\vec{n}, \psi) = e^{i\vec{s}\cdot\vec{n}\psi}$, where $\vec{s}$ is the spin of the particle ($\vec{s} = \vec{\sigma}/2$ for qubits), $\vec{n} = (\cos\phi\sin\theta, \sin\phi\sin\theta, \cos\phi)$, $\theta \in [0, \pi]$ and $\phi, \psi \in [0, 2\pi]$ [35].

Appendix C: Noise deconvolution for qubits

In this appendix we derive the noise deconvolution formula for qubits. Let ρ be a quantum state, and $\mathcal{N}$ a noise channel admitting an inverse map $\mathcal{N}^{-1}$, and $\hat{\mathcal{N}}^{-1}$ its adjoint map. Then, using Eq. (18) in (16), yields

$$
\begin{aligned}
\langle O \rangle &= \sum_\alpha \frac{1}{3}\operatorname{Tr}\left[\hat{\mathcal{N}}^{-1}\left(\mathbb{E}(O)[\sigma_\alpha]\right)\mathcal{N}(\rho)\right] \\
&= \sum_\alpha \frac{1}{3}\operatorname{Tr}\left[\left(\frac{3}{2}\operatorname{Tr}[O\sigma_\alpha]\hat{\mathcal{N}}^{-1}(\sigma_\alpha) \right.\right. \\
&\quad\left.\left. + \frac{1}{2}\operatorname{Tr}[O]\hat{\mathcal{N}}^{-1}(\mathbb{1})\right)\mathcal{N}(\rho)\right]
\end{aligned}
\tag{C.1}
$$

$$= \sum_\alpha \frac{1}{3}\left(\frac{3}{2}\operatorname{Tr}[O\sigma_\alpha]\langle\hat{\mathcal{N}}^{-1}(\sigma_\alpha)\rangle_{\mathcal{N}(\rho)}\right.$$

$$\left. + \frac{1}{2}\operatorname{Tr}[O]\underbrace{\operatorname{Tr}[\hat{\mathcal{N}}^{-1}(\mathbb{1})\mathcal{N}(\rho)]}_{=\operatorname{Tr}[\mathbb{1}\rho]=1}\right)$$

$$= \frac{1}{2}\operatorname{Tr}[O] + \frac{1}{2}\sum_{\alpha=x,y,z}\operatorname{Tr}[O\sigma_\alpha]\langle\hat{\mathcal{N}}^{-1}(\sigma_\alpha)\rangle_{\mathcal{N}(\rho)}.$$

Equation (C.1) lets us deconvolve the effect of noise by evaluating the expectation value of the noise-inverted Pauli matrices σ_α on the noisy state $\mathcal{N}(\rho)$. In particular, note that the formula remains valid whether the noise is *unital*—that is, $\mathcal{N}(\mathbb{1}) = \mathcal{N}^{-1}(\mathbb{1}) = \mathbb{1}$—or not. In fact, in the second line we can always move the adjoint inverse noise $\hat{\mathcal{N}}^{-1}$ from the identity to the noisy state $\mathcal{N}(\rho)$, thus obtaining $\operatorname{Tr}[\hat{\mathcal{N}}^{-1}(\mathbb{1})\mathcal{N}(\rho)] = \operatorname{Tr}[\mathbb{1}\mathcal{N}^{-1}(\mathcal{N}(\rho))] = \operatorname{Tr}[\rho] = 1$.

Appendix D: Inverse maps of noise channels

In this appendix we explicitly calculate the inverse map of the noise channels discussed in the main text.

D.1 Bit-flip, phase-flip, and bit-phase-flip channels

In the following we focus on the bit-flip channel, but the calculations are identical for the phase-flip and bit-phase-flip channels. The bit-flip channel is described by Kraus operators $A_0 = \sqrt{1-p}\mathbb{1}$ and $A_1 = \sqrt{p}\sigma_x$ so that its action is given by

$$\mathcal{N}_x(\rho) = (1-p)\rho + p\sigma_x\rho\sigma_x. \tag{D.1}$$

The Pauli Transfer Matrix Γ_x is defined as

$$(\Gamma_x)_{ij} = \frac{1}{2}\operatorname{Tr}[\sigma_i\mathcal{N}_x(\sigma_j)]. \tag{D.2}$$

By straightforward calculation one obtains

$$(\Gamma_x)_{11} = (1-p) + p = 1,$$
$$(\Gamma_x)_{22} = (1-p) - p = 1 - 2p,$$
$$(\Gamma_x)_{33} = (1-p) - p = 1 - 2p,$$
$$(\Gamma_x)_{ij} = 0, \quad \text{for } i \neq j$$

thus yielding

$$\Gamma_x = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & (1-2p) & 0 \\ 0 & 0 & 0 & (1-2p) \end{bmatrix}, \tag{D.3}$$

whose inverse is trivially

$$\Gamma_x^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{(2p-1)} & 0 \\ 0 & 0 & 0 & \frac{1}{(2p-1)} \end{bmatrix}. \tag{D.4}$$

The eigenvectors of such Pauli Transfer Matrix are clearly the Pauli matrices $\{|\mathbb{1}\rangle\rangle, |\sigma_x\rangle\rangle, |\sigma_y\rangle\rangle, |\sigma_z\rangle\rangle\}$ with eigenvalues $\boldsymbol{\lambda} = (1, 1, 1/(1-2p), 1/(1-2p))$.

The operator sum representation of $\mathcal{N}_x^{-1}$ can be reconstructed by noticing that the map

$$\mathcal{E}(O) = \sum_{j=0}^{3} \beta_j \sigma_j O \sigma_j, \tag{D.5}$$

has also the Pauli matrices as eigenvectors, but with eigenvalues $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_3)$. Since two maps are equals if they have the same action on a basis, then we can find the operator-sum representation of $\mathcal{N}_x^{-1}$ by finding those β_j such that $\boldsymbol{\lambda} = \boldsymbol{\beta}$. If we can find such mapping, then plugging those value in (D.5), we recover the operator sum of the inverse map.

The PTM matrix $\Gamma_{\mathcal{E}}$ of $\mathcal{E}$ amounts to

$$\Gamma_{\mathcal{E}} = \mathrm{diag}(\beta_0 + \beta_1 + \beta_2 + \beta_3,$$
$$\beta_0 + \beta_1 - \beta_2 - \beta_3,$$
$$\beta_0 - \beta_1 + \beta_2 - \beta_3,$$
$$\beta_0 - \beta_1 - \beta_2 + \beta_3).$$

The equality $\Gamma_x^{-1} = \Gamma_{\mathcal{E}}$ correspond to the system of equations

$$\begin{cases} 1 = \beta_0 + \beta_1 + \beta_2 + \beta_3, \\ 1 = \beta_0 + \beta_1 - \beta_2 - \beta_3, \\ \frac{1}{1-2p} = \beta_0 - \beta_1 + \beta_2 - \beta_3, \\ \frac{1}{1-2p} = \beta_0 - \beta_1 - \beta_2 + \beta_3, \end{cases} \tag{D.6}$$

where the first equation is the trace-preserving condition, dictated by the fact that the direct map is TP, and so the inverse map has to be. This condition is also evident from the expression of Γ_x^{-1} and $\Gamma_{\mathcal{E}}$, since the first row is of the form $(1, 0, 0, 0)$. The system of equations (D.6) has solutions

$$\beta_0 = \frac{1-p}{1-2p},$$
$$\beta_1 = -\frac{p}{1-2p},$$
$$\beta_2 = \beta_3 = 0$$

and substituting these values in Eq. (D.5) leads to the desired operator-sum representation

$$\mathcal{N}_x^{-1}(O) = \frac{1-p}{1-2p}O - \frac{p}{1-2p}\sigma_x O \sigma_x. \tag{D.7}$$

Similarly, the same procedure can be carried out for the dephasing (generated by σ_z) and bit-phase-flip channel (generated by σ_y), leading to

$$\mathcal{N}_z^{-1}(O) = \frac{1-p}{1-2p}O - \frac{p}{1-2p}\sigma_z O \sigma_z, \tag{D.8}$$

$$\mathcal{N}_y^{-1}(O) = \frac{1-p}{1-2p}O - \frac{p}{1-2p}\sigma_y O \sigma_y. \tag{D.9}$$

Note that for all these three cases the adjoint channels are equal to the direct ones, i.e. $\hat{\mathcal{N}}^{-1} = \mathcal{N}^{-1}$, since the generating operators are all Hermitian (see Appendix E for a case where this is not true).

We now proceed evaluating the explicit form of the deconvolution formula. Let $\beta \in \{x, y, z\}$ index one of the noise channels $\mathcal{N}_\beta \in \{\mathcal{N}_x, \mathcal{N}_y, \mathcal{N}_z\}$, the action of the inverse map on a Pauli matrix σ_α amounts to

$$\begin{aligned}
\mathcal{N}_\beta^{-1}(\sigma_\alpha) &= \frac{1}{1-2p}\big((1-p)\sigma_\alpha - p\sigma_\beta\sigma_\alpha\sigma_\beta\big) \\
&= \frac{1-2\delta_{\alpha\beta}p}{1-2p}\sigma_\alpha,
\end{aligned}$$

where in the second line we made use of the fact that $\sigma_\beta\sigma_\alpha\sigma_\beta = (2\delta_{\alpha\beta}-1)\sigma_\alpha$. Substituting this in Eq. (C.1), one obtains

$$\langle O \rangle_\beta = \frac{1}{2}\mathrm{Tr}[O] + \frac{1}{2(1-2p)}\sum_{\alpha=x,y,z}\mathrm{Tr}[O\sigma_\alpha](1-2\delta_{\alpha\beta}p)\langle\sigma_\alpha\rangle_{\mathcal{N}_\beta(\rho)},$$

where the subscript β in $\langle O \rangle_\beta$ is just used to denote that we are deconvolving with respect to noise $\mathcal{N}_\beta$, but remember that it correspond to the mitigated noise-free result.

Clearly, when the observable to be measured is itself a Pauli matrix $O = \sigma_\gamma$, this further simplifies to

$$\begin{aligned}
\langle \sigma_\gamma \rangle_\beta &= \frac{1}{2(1-2p)}\sum_{\alpha=x,y,z}\underbrace{\mathrm{Tr}[\sigma_\gamma\sigma_\alpha]}_{=2\delta_{\gamma\alpha}}(1-2\delta_{\alpha\beta}p)\langle\sigma_\alpha\rangle_{\mathcal{N}_\beta(\rho)} \\
&= \frac{1-2\delta_{\gamma\beta}p}{1-2p}\langle\sigma_\gamma\rangle_{\mathcal{N}_\beta(\rho)}.
\end{aligned}$$

D.2 Depolarizing channel
The Depolarizing channel is represented by the map

$$\mathcal{N}_{\mathrm{dep}}(\rho) = \left(1 - \frac{3p}{4}\right)\rho + \frac{p}{4}(\sigma_x\rho\sigma_x + \sigma_y\rho\sigma_y + \sigma_z\rho\sigma_z),$$

having Kraus operators $\{A_0 = \sqrt{1-3p/4}\,\mathbb{1}, A_1 = \sqrt{p}\sigma_x/2, A_2 = \sqrt{p}\sigma_y/2, A_3 = \sqrt{p}\sigma_z/2\}$.

By straightforward calculation, the Pauli Transfer Matrix amounts to

$$
\Gamma_{\text{dep}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1-p & 0 & 0 \\ 0 & 0 & 1-p & 0 \\ 0 & 0 & 0 & 1-p \end{bmatrix},
\tag{D.10}
$$

with inverse

$$
\Gamma_{\text{dep}}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{1-p} & 0 & 0 \\ 0 & 0 & \frac{1}{1-p} & 0 \\ 0 & 0 & 0 & \frac{1}{1-p} \end{bmatrix},
\tag{D.11}
$$

Following the same procedure used for the bit-flip channel, one arrives at the system of equations

$$
\begin{cases}
1 = \beta_0 + \beta_1 + \beta_2 + \beta_3, \\
\frac{1}{1-p} = \beta_0 + \beta_1 - \beta_2 - \beta_3, \\
\frac{1}{1-p} = \beta_0 - \beta_1 + \beta_2 - \beta_3, \\
\frac{1}{1-p} = \beta_0 - \beta_1 - \beta_2 + \beta_3
\end{cases}
\tag{D.12}
$$

which has solutions $\beta_0 = (4-p)/4(1-p)$ and $\beta_1 = \beta_2 = \beta_3 = -p/4(1-p)$. Substituting these values in (D.5), and using the relation $2\operatorname{Tr}[O]\mathbb{1} = O + \sigma_x O \sigma_x + \sigma_y O \sigma_y + \sigma_z O \sigma_z$, one obtains

$$
\mathcal{N}_{\text{depol}}^{-1}(O) = \frac{1}{1-p}\left(O - \frac{p}{2}\operatorname{Tr}[O]\mathbb{1}\right).
\tag{D.13}
$$

Plugging this in the tomographic deconvolution formula (C.1), leads to:

$$
\langle O \rangle = \frac{1}{2}\operatorname{Tr}[O] + \frac{1}{2}\sum_\alpha \frac{\operatorname{Tr}[O\sigma_\alpha]}{1-p}\langle \sigma_\alpha \rangle_{\mathcal{N}_{\text{dep}}(\rho)},
\tag{D.14}
$$

from which it is clear that whenever a Pauli matrix is to be measured, $O = \sigma_k$, then the expectation values are contracted by a factor $1-p$, i.e. $\langle \sigma_k \rangle = \langle \sigma_k \rangle_{\text{dep}}/(1-p)$.

D.3 General Pauli channel

The most general channel involving only Pauli operators is the arbitrary Pauli channel, given by

$$
\mathcal{N}_{\boldsymbol{p}}(\rho) = p_0 O + p_x \sigma_x \rho \sigma_x + p_y \sigma_y \rho \sigma_y + p_z \sigma_z \rho \sigma_z
\tag{D.15}
$$

characterized by probabilities $\boldsymbol{p} = (p_0, p_x, p_y, p_z)$, with the trace-preserving condition implying $p_0 = 1 - p_x - p_y - p_z$. The PTM of this map is diagonal

$$
\begin{aligned}
\Gamma_{\boldsymbol{p}} = \operatorname{diag}(1, & p_0 + p_x - p_y - p_z, \\
& p_0 - p_x + p_y - p_z, \\
& p_0 - p_x - p_y + p_z),
\end{aligned}
\tag{D.16}
$$

and has trivial inverse

$$
\begin{aligned}
\Gamma_{\boldsymbol{p}}^{-1} = \mathrm{diag}\big(1, (p_0 + p_x - p_y - p_z)^{-1}, \\
(p_0 - p_x + p_y - p_z)^{-1}, \\
(p_0 - p_x - p_y + p_z)^{-1}\big).
\end{aligned}
\tag{D.17}
$$

Again, using the same procedure as before, one arrives at the system of equations:

$$
\begin{cases}
1 = \beta_0 + \beta_1 + \beta_2 + \beta_3, \\
\frac{1}{p_0 + p_x - p_y - p_z} = \beta_0 + \beta_1 - \beta_2 - \beta_3, \\
\frac{1}{p_0 - p_x + p_y - p_z} = \beta_0 - \beta_1 + \beta_2 - \beta_3, \\
\frac{1}{p_0 - p_x - p_y + p_z} = \beta_0 - \beta_1 - \beta_2 + \beta_3,
\end{cases}
\tag{D.18}
$$

whose solution is reported in Eq. (41) in the main text. The action of the inverse map on the Pauli matrix σ_x is

$$
\begin{aligned}
\mathcal{N}_{\boldsymbol{p}}^{-1}(\sigma_x) &= \beta_0 \sigma_x + \beta_1 \sigma_x \sigma_x \sigma_x + \beta_2 \sigma_y \sigma_x \sigma_y + \beta_3 \sigma_z \sigma_x \sigma_z \\
&= (\beta_0 + \beta_1 - \beta_2 - \beta_3)\sigma_x \\
&= \frac{1}{1 - 2(p_y + p_z)}\sigma_x,
\end{aligned}
$$

and a similar expression also hold for σ_y and σ_z, from which one obtains the deconvolution formulas in Eq. (42).

Appendix E: Amplitude damping

The amplitude damping channel is given by the map

$$
\begin{aligned}
\mathcal{N}_{\mathrm{AD}}(\rho) &= K_0 \rho K_0^\dagger + K_1 \rho K_1^\dagger, \\
K_0 &= \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix}, \qquad K_1 = \begin{bmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{bmatrix}.
\end{aligned}
\tag{E.1}
$$

Differently from all the other cases treated above, this channel is not generated by coupled sigma matrices, and in addition one of its generators is not Hermitian. This has two consequences: first, we cannot straightforwardly apply the same eigenvalue matching procedure used above, second one must consider the adjoint channel when deconvolving.

The PTM of the amplitude damping channel is

$$
\Gamma_{\mathrm{AD}} = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & \sqrt{1-p} & 0 & 0 \\
0 & 0 & \sqrt{1-p} & 0 \\
p & 0 & 0 & 1-p
\end{bmatrix}
\tag{E.2}
$$

whose inverse is

$$\Gamma_{\text{AD}}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{1-p}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{1-p}} & 0 \\ \frac{-p}{1-p} & 0 & 0 & \frac{1}{1-p} \end{bmatrix}. \tag{E.3}$$

In this case the eigenvalues of Γ_{AD} and Γ_{AD}^{-1} are not the Pauli matrices, and so we cannot use the eigenvalue matching with the general map in (2). However, the two PTMs have the same structure, so one may easily guess that the operator-sum representation of the two maps share the same operators, something that also always happened in all previous cases. Let us then suppose that the inverse map $\mathcal{N}_{\text{AD}}^{-1}$ has the form

$$\mathcal{N}_{\text{AD}}^{-1}(\cdot) = \tilde{K}_0 \cdot \tilde{K}_0^\dagger - \tilde{K}_1 \cdot \tilde{K}_1^\dagger \tag{E.4}$$

with $\tilde{K}_0 = |0\rangle\langle 0| + \kappa |1\rangle\langle 1|$, and $\tilde{K}_1 = \tau |0\rangle\langle 1|$, with κ, τ free parameters to be determined. This map has PTM

$$\Gamma(\kappa, \tau) = \begin{bmatrix} \frac{1+\kappa^2-\tau^2}{2} & 0 & 0 & 0 \\ 0 & \kappa & 0 & 0 \\ 0 & 0 & \kappa & 0 \\ \frac{1-\tau^2-\kappa^2}{2} & 0 & 0 & \frac{1+\tau^2+\kappa^2}{2} \end{bmatrix}, \tag{E.5}$$

and by requiring that $\Gamma(\kappa, \tau) = \Gamma_{\text{AD}}^{-1}$, we obtain

$$\kappa = \frac{1}{\sqrt{1-\gamma}}, \qquad \tau = \sqrt{\frac{\gamma}{1-\gamma}},$$

thus recovering the inverse map

$$\mathcal{N}_{\text{AD}}^{-1}(O) = \tilde{K}_0 O \tilde{K}_0^\dagger - \tilde{K}_1 O \tilde{K}_1^\dagger,$$
$$\tilde{K}_0 = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\sqrt{1-\gamma}} \end{bmatrix}, \qquad \tilde{K}_1 = \begin{bmatrix} 0 & \sqrt{\frac{\gamma}{1-\gamma}} \\ 0 & 0 \end{bmatrix}. \tag{E.6}$$

In order to evaluate the deconvolution formula, we first need to calculate the adjoint of the inverse channel. Be Φ a linear map, its adjoint $\hat{\Phi}$ is defined as the unique map satisfying the following relation

$$\langle A, \Phi(B) \rangle = \langle \hat{\Phi}(A), B \rangle, \tag{E.7}$$

where $\langle \cdot, \cdot \rangle$ denotes the Hilbert-Schmidt inner product $\langle A, B \rangle \equiv \text{Tr}[A^\dagger B]$. Let's consider a generic linear map of the form

$$\Phi(A) = \sum_k \alpha_k V_k A V_k^\dagger, \qquad \alpha_k \in \mathbb{R}, \tag{E.8}$$

which is, in general, neither CP nor TP, since we make no further hypothesis on α_k and V_k. By direct application of the definition of adjoint map, we obtain

$$
\begin{aligned}
\langle A, \Phi(B) \rangle &\equiv \mathrm{Tr}\big[A^\dagger \Phi(B)\big] = \mathrm{Tr}\bigg[A^\dagger \sum_k \alpha_k V_k B V_k^\dagger\bigg] \\
&= \mathrm{Tr}\bigg[\sum_k \alpha_k V_k^\dagger A^\dagger V_k B\bigg] \\
&= \mathrm{Tr}\bigg[\bigg(\sum_k \alpha_k V_k^\dagger A V_k\bigg)^\dagger B\bigg] \\
&= \bigg\langle \sum_k \alpha_k V_k^\dagger A V_k, B\bigg\rangle \\
&\Rightarrow \quad \hat{\Phi}(A) = \sum_k \alpha_k V_k^\dagger A V_k,
\end{aligned}
$$

where we used the linearity and cyclic property of the trace, as well as the fact that the coefficients are real, $\alpha_k^* = \alpha_k \in \mathbb{R}$. We see that for any map of the form (E.8), its adjoint is obtained by simply substituting the operators with their adjoint, i.e. $V_k \to V_k^\dagger$. If the map Φ leverages only Hermitian operators $V_k = V_k^\dagger$, as it happens with every Pauli noise channel, than the adjoint and the direct map of course coincides, $\hat{\Phi}(\cdot) = \Phi(\cdot)$. However, the Amplitude Channel uses non Hermitian generators V_k, thus has a non-trivial, yet simple, adjoint map.

Straightforward application of the deconvolution formula then leads to the deconvolved expectation values

$$
\begin{aligned}
\langle \sigma_x \rangle &= \frac{1}{\sqrt{1-\gamma}} \langle \sigma_x \rangle_{\mathcal{N}_{\mathrm{AD}}(\rho)}, \\
\langle \sigma_y \rangle &= \frac{1}{\sqrt{1-\gamma}} \langle \sigma_y \rangle_{\mathcal{N}_{\mathrm{AD}}(\rho)}, \\
\langle \sigma_z \rangle &= \frac{1}{1-\gamma}\big(\langle \sigma_z \rangle_{\mathcal{N}_{\mathrm{AD}}(\rho)} - \gamma\big).
\end{aligned}
\tag{E.9}
$$

Appendix F: 2-Kraus channel

The set of channels considered here is generated by two parametrized Kraus operators

$$
\mathcal{N}_{\mathrm{two}}(\rho) = \sum_{i=1,2} A_i \rho A_i^\dagger,
\tag{F.1}
$$

with $A_1 = \cos\alpha |0\rangle\langle 0| + \cos\beta |1\rangle\langle 1|$, and $A_2 = \sin\beta |0\rangle\langle 1| + \sin\alpha |1\rangle\langle 0|$. The PTM of this channel and its inverse are respectively

$$
\Gamma_{\mathrm{two}} = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & \cos(\alpha-\beta) & 0 & 0 \\
0 & 0 & \cos(\alpha+\beta) & 0 \\
\frac{\cos(2\alpha)-\cos(2\beta)}{2} & 0 & 0 & \frac{\cos(2\alpha)+\cos(2\beta)}{2}
\end{bmatrix},
\tag{F.2}
$$

$$\Gamma_{\text{two}}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{\cos(\alpha-\beta)} & 0 & 0 \\ 0 & 0 & \frac{1}{\cos(\alpha+\beta)} & 0 \\ \frac{\cos(2\beta)-\cos(2\alpha)}{\cos(2\alpha)+\cos(2\beta)} & 0 & 0 & \frac{2}{\cos(2\alpha)+\cos(2\beta)} \end{bmatrix}. \tag{F.3}$$

Using the trigonometric relation

$$\cos(2\alpha) + \cos(2\beta) = 2\cos\left(\frac{2\alpha - 2\beta}{2}\right)\cos\left(\frac{2\alpha + 2\beta}{2}\right)$$

$$= 2\cos(\alpha - \beta)\cos(\alpha + \beta)$$

we can rewrite the elements of Γ_{two}^{-1} as

$$\left(\Gamma_{\text{two}}^{-1}\right)_{11} = h_{\alpha\beta}\cos(\alpha + \beta),$$

$$\left(\Gamma_{\text{two}}^{-1}\right)_{22} = h_{\alpha\beta}\cos(\alpha - \beta),$$

$$\left(\Gamma_{\text{two}}^{-1}\right)_{33} = h_{\alpha\beta}^2 \frac{\cos(2\alpha) + \cos(2\beta)}{2},$$

$$\left(\Gamma_{\text{two}}^{-1}\right)_{30} = h_{\alpha\beta}\frac{\cos(2\beta) - \cos(2\alpha)}{2},$$

with $h_{\alpha\beta} = \frac{2}{\cos(2\alpha)+\cos(2\beta)}$. Written in this way, these matrix elements are very similar to those in the PTM of the direct channel Γ_{two}. The differences are in the presence of the pre-factor $h_{\alpha\beta}$, as well as in the signs of the angles in elements $(\Gamma_{\text{two}}^{-1})_{11}$ and $(\Gamma_{\text{two}}^{-1})_{22}$, and in the sign in the difference in element $(\Gamma_{\text{two}}^{-1})_{30}$. This suggests that the operator-sum representation of the inverse map can be obtained starting from the direct one with some small changes, as it happened with the amplitude damping channel. First of all, we can multiply the Kraus operators by $\sqrt{h_{\alpha\beta}}$ to introduce the pre-factor, then, to account for the difference in elements $(\Gamma_{\text{two}}^{-1})_{11}$ and $(\Gamma_{\text{two}}^{-1})_{22}$, we can subtract the two operators instead of summing them. At last, element $(\Gamma_{\text{two}}^{-1})_{30}$ can be fixed by changing $\alpha \leftrightarrow \beta$ in the first Kraus operator A_1. Incidentally, these changes also fix the $(\Gamma_{\text{two}}^{-1})_{33}$ element to the correct value. Eventually, making these changes leads to defining new operators

$$B_1 = \sqrt{h_{\alpha\beta}}\cos(\beta)|0\rangle\langle 0| + \sqrt{h_{\alpha,\beta}}\cos(\alpha)|1\rangle\langle 1|,$$

$$B_2 = \sqrt{h_{\alpha\beta}}\sin(\beta)|0\rangle\langle 1| + \sqrt{h_{\alpha,\beta}}\sin(\alpha)|1\rangle\langle 0|,$$

$$h_{\alpha\beta} = \frac{2}{\cos(2\alpha) + \cos(2\beta)},$$

to be used within the inverse map

$$\mathcal{N}_{\text{two}}^{-1}(\cdot) = B_1 \cdot B_1^{\dagger} - B_2 \cdot B_2^{\dagger}.$$

One can check that this map has the desired Pauli Transfer Matrix Γ_{two}^{-1}.

As with the amplitude damping case, one the generators (B_2) is not Hermitian, thus one must be careful in considering the adjoint inverse map when evaluating the deconvolved mean values. By explicit calculations the following holds:

$$\langle\sigma_x\rangle = \frac{1}{\cos(\alpha - \beta)}\langle\sigma_x\rangle_{\mathcal{N}_{\text{two}}(\rho)},$$

$$\langle \sigma_y \rangle = \frac{1}{\cos(\alpha + \beta)} \langle \sigma_y \rangle_{\mathcal{N}_{\text{two}}(\rho)}, \tag{F.4}$$

$$\langle \sigma_z \rangle = h_{\alpha\beta} \left(\cos^2(\beta) + \sin^2(\alpha) - 1 + \langle \sigma_z \rangle_{\mathcal{N}_{\text{two}}(\rho)} \right).$$

Abbreviations
CPTP, Completely Positive Trace Preserving; PTM, Pauli Transfer Matrix; AD, Amplitude Damping; GAD, Generalized Amplitude Damping.

Additional information
The experiments with Aspen-9 were performed between the end of October and start of November 2021. The calibration data used in this analysis were downloaded from https://qcs.rigetti.com/lattices on the 30 October 2021.

Availability of data and materials
The data that support the findings of this study are available from the corresponding author, S.M., upon reasonable request.

Declarations

Competing interests
The authors declare no competing interests.

Author contributions
S.M, L.M and C.M conceived the main idea. S.M performed the analysis and simulations. All three authors discussed the results and approved the manuscript.

Author details
[1]Dipartimento di Fisica, Università di Pavia, Via Bassi 6, I-27100, Pavia, Italy. [2]INFN Sezione di Pavia, Via Bassi 6, I-27100, Pavia, Italy. [3]CNR-INO, Largo E. Fermi 6, I-50125, Firenze, Italy.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References
1. Steane AM. Phys Rev Lett. 1996;77:793–7. https://doi.org/10.1103/PhysRevLett.77.793.
2. Shor PW. Phys Rev A. 1995;52:R2493–6. https://doi.org/10.1103/PhysRevA.52.R2493.
3. Knill E. Nature. 2005;434:39–44. https://doi.org/10.1038/nature03350.
4. Preskill J. Quantum. 2018;2:79. https://doi.org/10.22331/q-2018-08-06-79.
5. Suzuki Y, Endo S, Fujii K, Tokunaga Y. Quantum error mitigation for fault-tolerant quantum computing. 2021. arXiv:2010.03887 [quant-ph].
6. Cao N, Lin J, Kribs D, Poon Y-T, Zeng B, Laflamme R. NISQ: error correction, mitigation, and noise simulation. 2021. arXiv:2111.02345 [quant-ph].
7. Viola L, Lloyd S. Phys Rev A. 1998;58:2733–44. https://doi.org/10.1103/PhysRevA.58.2733.
8. Bravyi S, Sheldon S, Kandala A, Mckay DC, Gambetta JM. Phys Rev A. 2021;103:042605. https://doi.org/10.1103/PhysRevA.103.042605.
9. Mari A, Shammah N, Zeng WJ. Phys Rev A. 2021;104:052607. https://doi.org/10.1103/PhysRevA.104.052607.
10. Temme K, Bravyi S, Gambetta JM. Phys Rev Lett. 2017;119:180509. https://doi.org/10.1103/PhysRevLett.119.180509.
11. Kandala A, Temme K, Córcoles AD, Mezzacapo A, Chow JM, Gambetta JM. Nature. 2019;567(7749):491–5. https://doi.org/10.1038/s41586-019-1040-7.
12. Endo S, Benjamin SC, Li Y. Phys Rev X. 2018;8:031027. https://doi.org/10.1103/PhysRevX.8.031027.
13. Lowe A, Gordon MH, Czarnik P, Arrasmith A, Coles PJ, Cincio L. Phys Rev Res. 2021;3:033098. https://doi.org/10.1103/PhysRevResearch.3.033098.

14. Siddhu V. Phys Rev A. 2019;99:012342. https://doi.org/10.1103/PhysRevA.99.012342.
15. Nielsen MA, Chuang IL. Quantum computation and quantum information. 10th ed. Cambridge: Cambridge University Press; 2010.
16. Lautenbacher L, de Melo F, Bernardes NK. Approximating invertible maps by recovery channels: optimality and an application to non-Markovian dynamics. 2021. arXiv:2111.02975 [quant-ph].
17. Gilyén A, Lloyd S, Marvian I, Quek Y, Wilde MM. Quantum algorithm for Petz recovery channels and pretty good measurements. 2020. arXiv:2006.16924 [quant-ph].
18. Wilde MM. Quantum information theory. 2nd ed. Cambridge: Cambridge University Press; 2017. https://doi.org/10.1017/9781316809976.
19. Karimipour V, Benatti F, Floreanini R. Phys Rev A. 2020;101:032109. https://doi.org/10.1103/PhysRevA.101.032109.
20. Shahbeigi F, Sadri K, Moradi M, Życzkowski K, Karimipour V. J Phys A, Math Theor. 2021;54:345301. https://doi.org/10.1088/1751-8121/ac13db.
21. Aurell E, Zakrzewski J, Życzkowski K. J Phys A, Math Theor. 2015;48:38FT01. https://doi.org/10.1088/1751-8113/48/38/38ft01.
22. Bény C. Quantum Inf Process. 2017;17:26. https://doi.org/10.1007/s11128-017-1796-3.
23. Karalekas PJ, Tezak NA, Peterson EC, Ryan CA, da Silva MP, Smith RS. Quantum Sci Technol. 2020;5:024003. https://doi.org/10.1088/2058-9565/ab7559.
24. de Pillis J. Pac J Math. 1967;23:129–37.
25. Bourdon PS, Williams HT. Phys Rev A. 2004;69:022314. https://doi.org/10.1103/PhysRevA.69.022314.
26. Jiang J, Wang K, Wang X. Quantum. 2021;5:600. https://doi.org/10.22331/q-2021-12-07-600.
27. Roncallo S, Maccone L, Macchiavello C. Multi-qubit noise deconvolution and characterization. arXiv; 2022. https://doi.org/10.48550/ARXIV.2207.12386.
28. D'Ariano GM, Chiribella G, Perinotti P. Quantum theory from first principles: an informational approach. Cambridge: Cambridge University Press; 2017. https://doi.org/10.1017/9781107338340.
29. Beth Ruskai M, Szarek S, Werner E. Linear Algebra Appl. 2002;347:159–87. https://doi.org/10.1016/S0024-3795(01)00547-X.
30. King C, Ruskai MB. IEEE Trans Inf Theory. 2001;47:192–209. https://doi.org/10.1109/18.904522.
31. Greenbaum D. Introduction to quantum gate set tomography. 2015. arXiv:1509.02921 [quant-ph].
32. Mauro D'Ariano G, Paris MGA, Sacchi MF. Quantum tomography. In: Advances in imaging and electron physics. vol. 128. Amsterdam: Elsevier; 2003. p. 205–308.
33. D'Ariano GM, Maccone L, Paris MGA. Phys Lett A. 2000;276:25–30. https://doi.org/10.1016/S0375-9601(00)00660-5.
34. D'Ariano GM, Lo Presti P. Phys Rev Lett. 2001;86:4195–8. https://doi.org/10.1103/PhysRevLett.86.4195.
35. D'Ariano GM, Maccone L, Paini M. J Opt B, Quantum Semiclass Opt. 2003;5:77–84. https://doi.org/10.1088/1464-4266/5/1/311.
36. D'Ariano GM. Phys Lett A. 2000;268:151–7. https://doi.org/10.1016/S0375-9601(00)00164-X.
37. Paris M, Řeháček J, editors. Quantum state estimation. Lecture notes in physics. Berlin: Springer; 2004.
38. D'Ariano GM, Maccone L, Paris MGA. J Phys A, Math Gen. 2001;34:93–103. https://doi.org/10.1088/0305-4470/34/1/307.
39. Bisio A, Chiribella G, D'Ariano GM, Facchini S, Perinotti P. IEEE J Sel Top Quantum Electron. 2009;15:1646–60. https://doi.org/10.1109/JSTQE.2009.2029243.
40. D'Ariano GM, Lo Presti P. Phys Rev Lett. 2003;91:047902. https://doi.org/10.1103/PhysRevLett.91.047902.
41. D'Ariano GM, Maccone L, Presti PL. Phys Rev Lett. 2004;93:250407. https://doi.org/10.1103/PhysRevLett.93.250407.
42. Vovrosh J, Khosla KE, Greenaway S, Self C, Kim MS, Knolle J. Phys Rev E. 2021;104:035309. https://doi.org/10.1103/PhysRevE.104.035309.
43. Wallman JJ, Emerson J. Phys Rev A. 2016;94:052325. https://doi.org/10.1103/PhysRevA.94.052325.
44. Hashim A, Naik RK, Morvan A, Ville J-L, Mitchell B, Kreikebaum JM, Davis M, Smith E, Iancu C, O'Brien KP, Hincks I, Wallman JJ, Emerson J, Siddiqi I. Randomized compiling for scalable quantum computing on a noisy superconducting quantum processor. 2021. arXiv:2010.00215 [quant-ph].
45. Emerson J, Alicki R, Życzkowski K. J Opt B, Quantum Semiclass Opt. 2005;7:S347–52. https://doi.org/10.1088/1464-4266/7/10/021.
46. Ware M, Ribeill G, Ristè D, Ryan CA, Johnson B, da Silva MP. Phys Rev A. 2021;103:042604. https://doi.org/10.1103/physreva.103.042604.
47. Ville J-L, Morvan A, Hashim A, Naik RK, Lu M, Mitchell B, Kreikebaum J-M, O'Brien KP, Wallman JJ, Hincks I, Emerson J, Smith E, Younis E, Iancu C, Santiago DI, Siddiqi I. Leveraging randomized compiling for the QITE algorithm. 2021. arXiv:2104.08785 [quant-ph].
48. D'Ariano GM, Sacchi MF. Phys Lett A. 2010;374:713–24. https://doi.org/10.1016/j.physleta.2009.11.081.
49. Huang H-Y, Kueng R, Preskill J. Nat Phys. 2020;16:1050–7. https://doi.org/10.1038/s41567-020-0932-7.
50. Cafaro C, Mancini S. Phys Rev A. 2010;82:012306. https://doi.org/10.1103/PhysRevA.82.012306.
51. Cafaro C, van Loock P. Phys Rev A. 2014;89:022316. https://doi.org/10.1103/PhysRevA.89.022316.
52. Smith RS, Curtis MJ, Zeng WJ. A practical quantum instruction set architecture. 2016. arXiv:1608.03355 [quant-ph].