IEEE Transactions on Quantum Engineering

Received January 15, 2021; revised June 24, 2021; accepted October 18, 2021; date of publication October 21, 2021; date of current version December 7, 2021.

Digital Object Identifier 10.1109/TQE.2021.3121797

New Single-Preparation Methods for Unsupervised Quantum Machine Learning Problems

YANNICK DEVILLE[1] (Member, IEEE), AND ALAIN DEVILLE[2]

[1]Institut de Recherche en Astrophysique et Planétologie, Observatoire Midi-Pyrénées, Centre National d'études Spatiales, Centre National de la Recherche Scientifique, Université Paul Sabatier, Université de Toulouse, F-31400 Toulouse, France

[2]Institut Matériaux Microélectronique Nanosciences de Provence (UMR 7334), Centre National de la Recherche Scientifique, Aix-Marseille Université, F-13397 Marseille, France

Corresponding author: Yannick Deville (yannick.deville@irap.omp.eu).

ABSTRACT The term "machine learning" especially refers to algorithms that derive mappings, i.e., input–output transforms, by using numerical data that provide information about considered transforms. These transforms appear in many problems related to classification/clustering, regression, system identification, system inversion, and input signal restoration/separation. We here analyze the connections between all these problems in the classical and quantum frameworks. We then focus on their most challenging versions, involving quantum data and/or quantum processing means, and unsupervised, i.e., blind, learning. We consider the general single-preparation quantum information processing (SIPQIP) framework that we have recently proposed. It involves methods that can work with a single instance of each (unknown) state, whereas usual methods are quite cumbersome because they have to very accurately create many copies of each (known) state. In our previous papers, we applied the SIPQIP approach to only one task [blind quantum process tomography (BQPT)], but it opens the way to a large range of other types of methods. In this article, we, therefore, propose various new SIPQIP methods that efficiently perform tasks related to system identification (blind Hamiltonian parameter estimation (BHPE), blind quantum channel identification/estimation, and blind phase estimation), system inversion and state estimation (blind quantum source separation (BQSS), blind quantum entangled state restoration (BQSR), and blind quantum channel equalization), and classification. Numerical tests show that our SIPQIP framework, moreover, yields much more accurate estimation than the usual multiple-preparation approach. Our methods are especially useful in a *quantum computer*, which we propose to more briefly call a "quamputer": BQPT and BHPE simplify the characterization of quamputer gates; BQSS and BQSR yield quantum gates that may be used to compensate for the nonidealities that alter states stored in quantum registers and open the way to very general self-adaptive quantum gates.

INDEX TERMS Blind Hamiltonian parameter estimation (BHPE), blind quantum process tomography (BQPT), blind quantum source separation (BQSS), blind quantum state restoration (BQSR), blind/unsupervised quantum machine learning, Heisenberg exchange coupling, quantum classification, quantum computer (quamputer), single-preparation quantum information processing (SIPQIP), spin-based qubits.

I. INTRODUCTION

Classical machine learning is currently a booming field [1], and various quantum machine learning extensions are also being considered [2]–[5]. The processing tasks that involve data-driven learning include not only widespread classification/clustering [1], [3], [6]–[10] and regression [6], [7], [9] but also especially: 1) classical system identification [11]–[13] and its quantum extension, called (nonblind [14]–[23] or blind [24]–[26]) quantum process tomography (QPT)[1]; 2) system inversion and signal restoration; and 3) blind source separation (BSS), e.g., based on independent component analysis (ICA) [30]–[36] (with a close connection with principal component analysis (PCA) [37]–[39]) and quantum

[1]Methods based on machine learning with neural networks have also been proposed for a partly related task, namely, quantum *state* tomography [27]–[29].

TABLE I Application-Dependent Terminology of Classical and Quantum Machine Learning Tasks (i.e., Data-Driven Learning/Adaptation Methods) Related to System Identification, System Inversion, Signal Restoration, and Source Separation (Apart From Classification/Clustering and Regression, Addressed in Section V). This Yields Nonblind and Blind (i.e., Supervised and Unsupervised) Variants. The Sections of This Article Mainly Dealing With the Blind (i.e., Unsupervised) Quantum Single-Preparation Versions of These Methods Are Mentioned

	system identification	system inversion, signal restoration, source separation
classical signals	• channel identification, channel estimation (communication) • impulse response or transfer function estimation (acoustics...)	• channel equalization (communication) • deconvolution (image, seismology...) • dereverberation (acoustics) • cocktail party processing (audio) • deblurring (image)
quantum states	• process tomography (Appendix C) • Hamiltonian estimation (Section III-A) • channel estimation (Section III-B) • phase estimation (Section III-B)	• source separation (Section IV-A) • state restoration (Section IV-B) • channel equalization (Section IV-C)

extensions of BSS/ICA [40]–[45]. Moreover, in various application fields, these tasks are given different names (see the summary in Table 1), especially channel identification or channel estimation [12], (channel) equalization [12], [46], dereverbation [47], deconvolution [48], deblurring [48], or cocktail party problem [49].

Beyond their apparent diversity, the above data processing tasks share major features that are analyzed in detail in Appendix A and that may be summarized as follows. They involve mappings from input data to output data, and these mappings are derived from a set of known values of these input and/or output quantities, depending whether that learning is performed in the so-called supervised or unsupervised (i.e., nonblind or blind) modes, whose definitions depend on the considered task and are also analyzed in Appendix A. These approaches are developed in order to characterize the mapping performed by a given natural or artificial system and/or in order to build an artificial system that performs a mapping (i.e., a data transformation) suited to the considered application.

In this article, we investigate a variety of the above-defined data processing tasks, and we focus on advanced configurations from the following points of view. First, we only consider a quantum framework in terms of the nature of the data to be processed and/or of the means used to process them. Second, we almost only address unsupervised learning, which is more challenging than supervised learning because it consists of learning mappings without known values (but with a few known properties) for the input or output of that mapping. The overview of classical and quantum machine learning provided in Appendix A includes references to the currently quite limited set of works from the literature, which is dedicated to that quantum and unsupervised learning framework that we tackle in this article. Moreover, we here proceed beyond that framework, by adding another feature: We focus on what we call "single-preparation operation," as defined in the following. *So, to summarize, the investigations reported in this article fall in the pre-existing field of unsupervised quantum machine learning, and our original contribution in this article consists of various processing methods based on our single-preparation concept, to be contrasted with the "multiple-preparation methods" defined in the following, which are usually employed in the literature.*

More precisely, our single-preparation concept (detailed in Section II) may be defined as follows. Various quantum machine learning methods from the literature, e.g., intended for system identification (i.e., say, QPT) or system inversion, use multiple-preparation approaches in the sense that, for each quantum state value that they consider, they estimate the probabilities of corresponding measurement outcomes by using the sample frequencies of these outcomes over a set of measurements, which requires a set of copies of the considered quantum state, to perform one measurement for each copy (see details in Appendix B). In contrast, we have very recently introduced a statistical approach, which yields much higher flexibility since it avoids the burden of very accurately preparing many ideally identical copies of the same known state, by allowing one to replace these copies by a set of states whose values are possibly different and unknown but only requested to belong to a general known class [26], [50]. The definition of this single-preparation concept is summarized in Section II-A. This concept is quite general, but, in [26] and [50], we only detailed its application to a single data processing task, namely, single-preparation blind (i.e., unsupervised) QPT, which is summarized in Appendix C. In this article, we aim at showing how this single-preparation processing concept may be applied to a variety of other quantum information processing (QIP) tasks of interest, thus yielding a general "single-preparation quantum information processing" (SIPQIP) framework.

The terminology used in this article deserves the following comments. Quantum mechanics (QM) considers that an isolated quantum system may be either in a pure state—the result of some preparation—described by a ket with deterministic coefficients (in the Schrödinger picture), or more generally in a state called a mixed state or a statistical mixture, usually described by a density operator. When developing our methods, first in the blind quantum source separation (BQSS) field and then in the blind quantum process tomography (BQPT) field, we were led to distinguish between what we hereafter call a "deterministic-coefficient pure state" (the usual pure state of QM) and a "random-coefficient pure state," described by a ket with random-valued coefficients when developed over an orthonormal basis of fixed vectors. The relationships between a random-coefficient pure state and a mixed state have been analyzed in [51]. Deterministic-coefficient pure states may also be considered as a specific subset of random-coefficient pure states, corresponding to

the case when the random variables that define the ket coefficients of random-coefficient pure states reduce to fixed values, i.e., with "no uncertainty." In the context of BQSS or BQPT, depending on the considered method, the system of interest is initialized either in a deterministic-coefficient or in a random-coefficient pure state. Most of the methods proposed in this article are based on qubits initialized with random-coefficient pure states. Such an initialization is also called a state preparation hereafter.

The remainder of this article is organized as follows. In Section II, we define the single-preparation quantum processing concept, which is our original, very recent, and general feature then used in all the processing methods proposed for the first time in this article. These methods deal with various problems related to quantum system identification (see Section III), quantum system inversion and state restoration (see Section IV), and quantum classification (see Section V). Finally, Section VI contains conclusions about the processing tasks addressed in this article and a discussion of potential extensions of the proposed methods to other QIP problems.

II. SINGLE-PREPARATION QIP

The standard use of quantum state preparation and measurements in QIP [14] was defined in [26] and [50]. In Appendix B, we summarize its main features required in this article, and we introduce the corresponding notations, which are used hereafter. As explained in Appendix B, that approach is called "multiple-preparation QIP" because it requires many copies of the *same* quantum state. This is, therefore, constraining, especially in the framework of *blind* QIP, where the processing methods should operate with unknown values of some quantum states (e.g., unknown inputs of the process to be identified with QPT). Indeed, being able to operate without requiring known values of quantum states in blind methods is attractive, but then requesting many copies of each such state to be available is still a limitation, because it still requires some form of control of these states. We would like to avoid that limitation, in order to simplify the practical operation of the considered methods and to make them "blinder." We hereafter provide a solution to this problem.

A. SINGLE-PREPARATION QIP BASED ON PROBABILITY EXPECTATIONS

The description of "multiple-preparation QIP" in Appendix B was provided for an arbitrarily selected deterministic-coefficient pure quantum state. We had to extend that framework to *random-coefficient* pure quantum states when we developed our first class of BQSS methods (see, e.g., [40]–[42] and [51]) and associated BQPT methods (see, e.g., [24], [25], and [51]). The concept of random-coefficient pure states may be summarized as follows (see [51] for more details). Whereas the coefficients c_k in (27) are fixed parameters for deterministic-coefficient pure states, they become complex-valued random variables for random-coefficient pure states. Therefore, the probabilities $P(A_k)$ in (28) also become random variables for random-coefficient pure states.

We here analyze how to estimate the expectations of the above-defined random variables $P(A_k)$, denoted as $E\{P(A_k)\}$, which will then be repeatedly used in the original methods proposed in this article. This may be done either with a natural extension of the approach defined in Appendix B (see details in [26] and [50]) or, preferably, by using the original and more flexible, single-preparation, approach that we introduced in [26] and [50]. We applied that approach in detail to a single QIP task (namely, BQPT) in [26] and [50], whereas, in this article, we show that it also applies to a wide range of other QIP tasks. For the sake of clarity, we here first summarize the principle of this single-preparation approach. In practice, to estimate any expectation $E\{P(A_k)\}$, one replaces the expectation operator $E\{.\}$ by a sample mean, i.e., essentially by "a *sum*." More precisely, one employs the normalized sum of values that reads

$$E'\{P(A_k)\} = \frac{\sum_{n=1}^{N} P(A_k, n)}{N} \tag{1}$$

where one uses only a *finite* number N of states $|\psi(n)\rangle$ and $P(A_k, n)$ are the associated values of the above-defined probability $P(A_k)$.

Moreover, in (1), each probability $P(A_k, n)$ is replaced by a sample frequency, i.e., essentially by "a *sum*." More precisely, one uses a sum of 1 and 0, depending whether the considered event occurs or not for each trial defined by a preparation of the considered quantum state and by a measurement of the associated spin components. This summation is here again followed by a normalization, by the total number of trials. Each probability $P(A_k, n)$ in (1) is, thus, replaced by the following approximate value:

$$P'(A_k, n, K) = \frac{\mathcal{N}(A_k, n, K)}{K} \tag{2}$$

where $\mathcal{N}(A_k, n, K)$ is the number of occurrences of event A_k for the state $|\psi(n)\rangle$ when performing measurements for a set of K copies of that state $|\psi(n)\rangle$.

When combining the above two approximations, $E\{P(A_k)\}$ is estimated by a (normalized) "sum of sums," which reads

$$E''\{P(A_k)\} = \frac{\sum_{n=1}^{N} \mathcal{N}(A_k, n, K)}{NK}. \tag{3}$$

This two-level sum may then be reinterpreted as a single global sum, as follows. In (3), the quantity $\sum_{n=1}^{N} \mathcal{N}(A_k, n, K)$ is equal to the number of occurrences of event A_k for the complete considered set of $L = NK$ measurements. That number of occurrences is hereafter denoted as $\mathcal{N}(A_k, L)$. The quantity $E''\{P(A_k)\}$ in (3) is, thus, the relative frequency of occurrence of event A_k over these L measurements. These measurements may be called "trials," using standard probabilistic terms [52]. The quantity $E''\{P(A_k)\}$ in (3) may also be expressed as

$$E''\{P(A_k)\} = \frac{\mathcal{N}(A_k, L)}{L} \tag{4}$$

$$= \frac{\sum_{\ell=1}^{L} \mathbb{1}(A_k, \ell)}{L} \tag{5}$$

where $\mathbb{1}(A_k, \ell)$ is the value of the indicator function of event A_k for trial ℓ. That function is equal to 1 if A_k occurs during trial ℓ, and 0 otherwise. When using (5), one now considers the $L = NK$ trials as organized as a single series, with trials indexed by ℓ. One, thus, fuses the above-defined two levels of summation into a single one. In that approach, one, therefore, does not take into account that, in this series, each block of K consecutive trials uses the same state $|\psi(n)\rangle$. In [26] and [50], we proved that the number K of used copies of each state $|\psi(n)\rangle$ may be freely decreased, and even set to one, while keeping the same total number L of trials. Setting $K = 1$, moreover, yields better performance [26], [50]. One, thus, obtains the proposed single-preparation framework, referred to as SIPQIP. Moreover, the proposed estimator (5) of $E\{P(A_k)\}$ is attractive because it is asymptotically efficient in the conditions defined in [26] and [50].

It should be clear that this procedure for estimating $E\{P(A_k)\}$, and hence the resulting SIPQIP methods, can be freely used with either one instance or several (e.g., many) copies per state. This SIPQIP terminology, therefore, means that these methods *allow* one to use a single instance of each state. In contrast, so-called multiple-preparation QIP methods can achieve good performance only by *forcing* one to use many state copies.

B. SINGLE-PREPARATION QIP BASED ON SAMPLE MEANS OF PROBABILITIES

As explained in Section II-A, the framework introduced in that section is intended for a formalism based on *random-coefficient* pure states, which we use in most of this article. In addition, we employ the more standard formalism of *deterministic-coefficient* pure states in Section V. In that section, the considered QIP task eventually boils down to estimating the mean, over a finite number (i.e., the sample mean) of deterministic-coefficient pure states, of the (hence deterministic) probabilities, respectively, associated with each of these states. Although this framework is conceptually different from the one of Section II-A, it eventually yields the same implementation as will now be shown. Here again, each of the considered probabilities is replaced by (2) in practice (whereas the corresponding theoretical probability value is the limit of (2) when K tends to infinity, in the frequentist approach to probabilities). The difference with respect to Section II-A then appears when combining all these elementary probability estimates (2). This difference occurs because we here directly aim at handling a *finite* number of quantum states, so that we directly use the quantity in (1) and its estimate (4), instead of first considering the corresponding expectation $E\{P(A_k)\}$. The remainder of the analysis of Section II-A then also applies to the framework considered here. Our SIPQIP concept, therefore, also applies to this framework and, thus, here again yields the above-defined advantages.

III. QIP TASKS RELATED TO SYSTEM IDENTIFICATION

As explained in Section I and Appendix A, the quantum version of system identification is often referred to as QPT. It is of major importance, especially for characterizing the actual behavior of quantum gates (see, e.g., [14], [17], [18], [20], [21], and [23]), which are the building blocks of a *qua*ntum co*mputer*, which, by the way, we propose to more briefly call a "quamputer." We introduced blind and single-preparation extensions of QPT in our previous papers (see [50] for a partial version and [26] for complete extensions). The practical versions of these BQPT methods that we detailed are targeted at a device composed of two distinguishable [45] qubits implemented as electron spins 1/2, which are internally coupled according to the cylindrical-symmetry Heisenberg model.

BQPT is the only task related to system identification, for which we detailed SIPQIP solutions in our previous papers. In the present section, we introduce new SIPQIP methods for other QIP tasks also related to system identification. Since several of these and subsequent methods proposed in this article build upon our aforementioned single-preparation BQPT methods, we summarize the major features of the latter methods in Appendix C, for the sake of readability.

A. BLIND HAMILTONIAN PARAMETER ESTIMATION (BHPE)

1) PROPOSED METHOD

As shown in [26], the behavior of the device composed of two Heisenberg-coupled qubits that we mentioned above is primarily defined by its Hamiltonian. The associated process matrix M involved in BQPT [see also (13)–(17)] then follows when considering the evolution of the state of that system from a fixed time t_0 to a fixed time t. Therefore, beyond the estimation of the process matrix M, a related QIP task consists of estimating the primary unknown parameters of the Hamiltonian of the studied device, namely, the principal values J_{xy} and J_z of the exchange tensor (similar considerations are also provided in [53]). This type of task (for the parameters of this or other Hamiltonians) is called Hamiltonian parameter estimation (HPE) hereafter and also especially in [53]–[55]. Such parameter estimation problems are also addressed but often referred to as Hamiltonian identification, e.g., in [22], [56], [57], and partly [58]. To our knowledge, in the literature, this task has been studied only in the nonblind or "controlled" mode and/or using multiple preparations in approaches that are closely connected with the conventional QPT [58] or that are based on specific protocols. These protocols include periodical sampling (hence with a potentially quite high total number of required state preparations) [22], [53], [57], use of a closed-loop [56] or optimal feedback [55] structure, or curve fitting with respect to the experimental results obtained for various angles of the magnetic field [54]. In contrast, we, hereafter, investigate a single-preparation and blind (without control) version of this HPE task, based on measurements along the Oz and Ox axes, which we did not address in our previous papers. This approach has direct

IEEE Transactions on Quantum Engineering

connections with the above-defined single-preparation and blind version of QPT. It yields the same type of attractive features as for QPT. It avoids the burden of very accurately and repeatedly preparing predefined states to estimate the unknown parameters of the considered Hamiltonian. Here again, we show how to develop such an extension for the specific class of Hamiltonians corresponding to Heisenberg coupling with unknown J_{xy} and J_z. However, this should again be considered only as an example, which the reader may then extend to other types of Hamiltonians. Similarly, various investigations in the literature considered specific parameterized Hamiltonians, with a limited number of unknown parameters, as the core of the proposed approaches or to illustrate them (see, e.g., [22] and [53], [54] and [56]–[58]). Other early works related to Hamiltonian identification and associated control tasks, e.g., include [59]–[61].

The BHPE method that we propose builds upon the BQPT algorithm summarized in Appendix C, but it requires subsequent developments for the following reason. As explained in Appendix C, the BQPT method is strongly connected with estimating the quantities $\exp[i\frac{J_{xy}(t-t_0)}{\hbar}]$ and $\exp[i\frac{J_z(t-t_0)}{2\hbar}]$, using some types of measurements. The estimation of these very quantities would define their phase arguments $\frac{J_{xy}(t-t_0)}{\hbar}$ and $\frac{J_z(t-t_0)}{2\hbar}$, only up to additive integer multiples of 2π, which would yield the indeterminacies of this estimation procedure from the point of view of BHPE. More precisely, using the data provided by the considered measurements, the above BQPT method yields the indeterminacies that consist of the additive constants $\widehat{k}_{xy}\pi$ and $2\widehat{k}_z\pi$ of (44) and (46). It, thus, does not provide a unique solution with respect to $\frac{\widehat{J}_{xy}\tau_1}{\hbar}$ and $\frac{\widehat{J}_z\tau_2}{\hbar}$ and, hence, $\widehat{J}_{xy}$ and $\widehat{J}_z$. Therefore, it does not solve the HPE problem considered here (related comments may be found in [53]). For instance, let us consider the test conditions defined in Appendix E, including the available prior knowledge about the range of values to which J_{xy} is guaranteed to belong. Then, a single run of our BQPT method yields 32 acceptable determinations of J_{xy} in that range and no means to know which of these numerous potential solutions corresponds to the actual value J_{xy}.

We here aim at developing a BHPE method that takes advantage of the above BQPT algorithm so as to estimate $\widehat{J}_{xy}$ and $\widehat{J}_z$ without indeterminacies. The trick that we propose to this end is based on estimating each of the parameters J_{xy} and J_z by using *two* values of the above-defined time interval $(t-t_0)$, instead of one in the fundamental principle of the above BQPT method. This trick also has relationships with the practical approach that we used in [26], for BQPT only: Starting from a basic BQPT method that uses a single value of $(t-t_0)$ and that, thus, yields some indeterminacies with respect to M, we then moved to a more advanced BQPT method, which uses several values of $(t-t_0)$ and, thus, avoids all indeterminacies with respect to M (this is the method summarized in Appendix C). However, for BQPT, we, thus, eventually used several values of $(t-t_0)$ for the complete practical procedure but only one value for each part of that BQPT method, e.g., associated with the phase factor involving one of the parameters J_{xy} and J_z (see (44) and (46), respectively). Instead, for BHPE, we here move to two values of $(t-t_0)$ per parameter J_{xy} and J_z. These values are, thus, exploited in a new way that we describe hereafter.

Let us first consider the estimation of J_{xy}. To this end, we use the procedure of the first part of the BQPT method of Appendix C. We apply it twice, with τ_1 of Appendix C successively replaced by two values denoted as τ_{11} and τ_{12}. We combine (41) and (44), with τ_1 replaced by τ_{11} and similarly with an additional index "1" for the other variables, whose values are specific to that first application of the procedure. This yields

$$\widehat{J}_{xy1} = J_{xy} + \frac{\hbar}{\tau_{11}}\left(\Delta_{Ed1} - \widehat{\Delta}_{Ed1} + \Delta k_{xy1}\pi\right) \tag{6}$$

with

$$\Delta k_{xy1} = \widehat{k}_{xy1} - k_{xy1}. \tag{7}$$

This shows that the procedure applied with the time interval τ_{11} yields a regular 1-D grid of possible estimates $\widehat{J}_{xy1}$ of J_{xy} (associated with the values of $\widehat{k}_{xy1}$), with a step equal to $\frac{\hbar\pi}{\tau_{11}}$. Similarly, the second application of that procedure, with a time interval τ_{12}, yields

$$\widehat{J}_{xy2} = J_{xy} + \frac{\hbar}{\tau_{12}}\left(\Delta_{Ed2} - \widehat{\Delta}_{Ed2} + \Delta k_{xy2}\pi\right) \tag{8}$$

with

$$\Delta k_{xy2} = \widehat{k}_{xy2} - k_{xy2}. \tag{9}$$

The corresponding estimates $\widehat{J}_{xy2}$ of J_{xy}, therefore, form a regular grid with a step equal to $\frac{\hbar\pi}{\tau_{12}}$.

We here aim at exploiting the differences between the above two[2] grids of values. As a preliminary stage, let us consider the ideal case, i.e., when

$$\widehat{\Delta}_{Ed1} = \Delta_{Ed1} \qquad \text{and} \qquad \widehat{\Delta}_{Ed2} = \Delta_{Ed2}. \tag{10}$$

Then, the above two grids share at least one value, equal to J_{xy} and obtained when $\widehat{k}_{xy1}$ and $\widehat{k}_{xy2}$ are, respectively, set to k_{xy1} and k_{xy2}, which results in $\Delta k_{xy1} = 0$ and $\Delta k_{xy2} = 0$. Moreover, let us consider the case when τ_{12}/τ_{11} is set to an irrational value. Then, the above grids only share the value J_{xy}. This occurs because (6) and (8) show that, when (10) is met, the values of $\widehat{k}_{xy1}$ and $\widehat{k}_{xy2}$ that are such that the corresponding estimates $\widehat{J}_{xy1}$ and $\widehat{J}_{xy2}$ are equal are those that meet $\Delta k_{xy1}/\tau_{11} = \Delta k_{xy2}/\tau_{12}$ so that, when Δk_{xy1} and Δk_{xy2} are nonzero, this requires τ_{12}/τ_{11} to be equal to the rational value $\Delta k_{xy2}/\Delta k_{xy1}$. So, when (10) is met and τ_{12}/τ_{11} is set to an irrational value, a simple criterion for determining J_{xy} is: It is the only value shared by the above two grids. This behavior has a relationship with the influence of the sampling period when sampling a sine wave, as, e.g., detailed in [62]. Oppenheim and Schafer [62], thus, also indirectly show that

[2]This approach might be further extended to more than two grids to make it more robust.

the above attractive behavior of our grids is obtained *only* if τ_{12}/τ_{11} is irrational.

The above criterion must, then, be modified when moving to practical situations, because the available estimates $\widehat{\Delta}_{Ed1}$ and $\widehat{\Delta}_{Ed2}$ are somewhat and independently shifted with respect to the corresponding actual values. Therefore, to estimate J_{xy}, instead of looking for values of both grids, which are identical in the ideal case, this here suggests to compare each value $\widehat{J}_{xy1}$ of the first grid to each value $\widehat{J}_{xy2}$ of the second grid in order to derive the couple of closest values. Moreover, in practical configurations, one usually has prior knowledge about a range of values to which J_{xy} and, hence, its relevant estimates are guaranteed to belong. This known range may be exploited in such a way that the values of $\widehat{J}_{xy1}$ and $\widehat{J}_{xy2}$, which are the closest to one another in this range, are also those which are the closest to J_{xy}: To this end, one may use the method detailed in Appendix D. From these two specific values, an estimate of J_{xy} is eventually derived as $\dfrac{\widehat{J}_{xy1}+\widehat{J}_{xy2}}{2}$.

Similarly, the parameter J_z is estimated by using the procedure of the second part of the BQPT method of Appendix C, based on (45) and (46). This procedure is here applied twice, with different values of the parameter τ_2 of Appendix C. The resulting method is described in Appendix D.

2) TEST RESULTS

The physical implementation of qubits is an emerging topic, which is beyond the scope of this article. We, therefore, assessed the performance of the proposed BHPE method by means of numerical tests performed with data derived from a software simulation of the considered configuration. Each elementary test consists of the following stages. We first create a set of N input states $|\psi(t_0)\rangle$. Each such state is obtained by randomly drawing its six parameters r_j, θ_j, and ϕ_j, with $j \in \{1, 2\}$ and then using (32)–(34) (the state (32) is defined by the above six parameters, but only the four parameters r_j and $\phi_j - \theta_j$ have a physical meaning). We then transfer these states $|\psi(t_0)\rangle$ through the quantum process to be identified (see details in [26]). This uses given values of the parameters of the considered Hamiltonian and, hence, of the matrix M of the studied quantum process. This yields the states $|\psi(t)\rangle$. More precisely, we eventually use simulated measurements of spin components associated with these states $|\psi(t)\rangle$. For measurements along the Oz-axis, this means that we use the model (35)–(37) with a given value of the parameter v, corresponding to the above values of the parameters of the considered Hamiltonian. For each of the N states $|\psi(t_0)\rangle$, corresponding to parameter values (r_1, r_2, Δ_I), (35)–(37), thus, yield the corresponding set of probability values $(p_{1zz}, p_{2zz}, p_{4zz})$, which are used as follows. We use K prepared copies of the considered state $|\psi(t_0)\rangle$ to simulate K random-valued two-qubit spin component measurements along the Oz-axis, drawn with the above probabilities $(p_{1zz}, p_{2zz}, p_{4zz})$. We then derive the sample frequencies of the results of these K measurements, which are estimates of p_{1zz}, p_{2zz}, and p_{4zz} for the considered state $|\psi(t_0)\rangle$ [see (2)]. Then, computing the averages of these K-preparation estimates over all N source vectors $|\psi(t_0)\rangle$ yields (NK)-preparation estimates of probability expectations $E\{p_{kzz}\}$ [see (3)]. Spin component measurements for the Ox-axis are handled similarly (with other state preparations), thus yielding estimates of probability expectations $E\{p_{kxx}\}$. Both types of estimates of probability expectations are then used by our BHPE method defined in Section III-A1 to derive the estimates $\widehat{J}_{xy}$ and $\widehat{J}_z$.

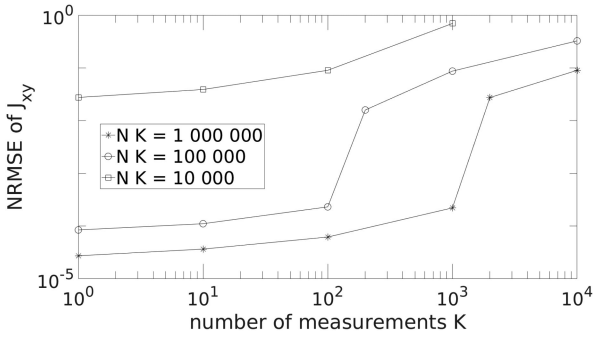

FIG. 1. NRMSE of estimation of parameter J_{xy} versus number K of preparations of each of the N used states.

In these tests, the above parameters N and K were varied as described further in this section, whereas the numerical values of the other parameters were fixed as explained in Appendix E. We, thus, used the same values for the parameters of the considered Hamiltonian in all tests. For each considered set of conditions defined by the values of N and K, we performed 100 above-defined elementary tests, with different sets of states $|\psi(t_0)\rangle$, in order to assess the statistical performance of the considered BHPE method over up to 100 estimations of the same set $\{J_{xy}, J_z\}$ of parameter values. More precisely, all 100 estimates of J_{xy} were real-valued and were kept. In contrast, for some test conditions, some estimates of J_z were complex-valued (because they were derived from trigonometric equations, where some *estimates* of sines or cosines may be situated out of the interval $[-1, 1]$). Since these false values can actually be detected and rejected in practice, the estimation performance for J_z was computed only over its real-valued estimates.

The considered performance criteria are defined as follows. Separately for each of the parameters J_{xy} and J_z, we computed the normalized root-mean-square error (NRMSE) of that parameter over all considered estimates, defined as the ratio of its root-mean-square error to its actual (positive) value. The values of these two performance criteria are shown in Figs. 1 and 2, where each plot corresponds to a fixed value of the product NK, i.e., of the complexity of the BHPE method in terms of the total number of state preparations. Each plot shows the variations of the considered performance criterion versus K, hence with N varied accordingly, to keep the considered fixed value of NK.

Figs. 1 and 2 first show that the proposed BHPE method is able to operate with a number K of preparations per state

IEEE Transactions on Quantum Engineering

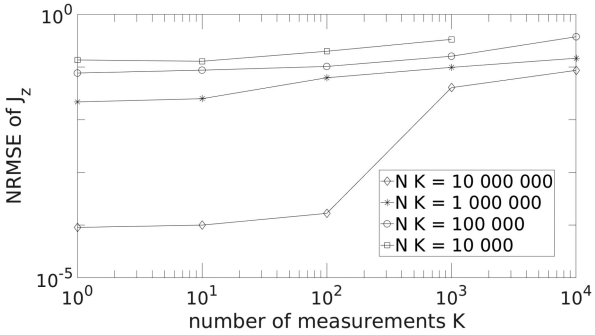

FIG. 2. NRMSE of estimation of parameter J_z versus number K of preparations of each of the N used states.

$|\psi(t_0)\rangle$ decreased down to one, as expected. Moreover, for a fixed value of NK, the errors decrease when K decreases, which is expected to be due to the fact that the number N of *different* used states, thus, increases, allowing the estimation method to better explore the statistics of the considered random process. The magnitude of the error reduction from the highest value of K down to $K = 1$ is often quite large, especially for J_{xy}, that is, between one and two orders of magnitude even when disregarding the "discontinuity" in some plots discussed hereafter. *This means that the proposed SIPQIP framework is then of high interest not only in terms of simplicity of operation of QIP methods, but also with respect to their accuracy.*

Moreover, some of the plots contain the aforementioned type of discontinuity. For example, in Fig. 1, the NRMSE of J_{xy} for the fixed value $NK = 100\,000$ abruptly decreases from around 2×10^{-2} when $K = 200$ to around 2×10^{-4} when $K = 100$. This behavior is normal: It is due to the intrinsically discontinuous nature of the specific type of estimation algorithm used here for J_{xy} (the same considerations apply to J_z, as confirmed by Fig. 2). More precisely, in conditions when J_{xy} is estimated with a low accuracy, the following phenomenon may occur for one or several runs of the estimation procedure: That procedure may select a false determination of the estimate of J_{xy}, that is, a value corresponding to false (i.e., nonzero) values of Δk_{xy1} in (6) and Δk_{xy2} in (8). The estimated value of J_{xy} is, thus, strongly shifted, because, e.g., the corresponding values on the first grid (6) are shifted by multiples of the step $\frac{\hbar\pi}{\tau_{11}}$, as explained above. For the numerical values considered here, the corresponding step for the determinations of $\widehat{J}_{xy}/k_B$ is $\frac{\hbar\pi}{\tau_{11}k_B} \simeq 0.048$ K, as compared to the actual value of J_{xy}/k_B equal to 0.3 K in these tests. Therefore, a shift equal to one step, i.e., obtained with $\Delta k_{xy1} = 1$, corresponds to a relative error for $\widehat{J}_{xy}/k_B$, and, hence, for $\widehat{J}_{xy}$, around 16% for the considered estimate of J_{xy}. The overall error for 100 estimates then depends on the number of runs, where such false determinations are selected, but as long as at least one of them is selected, the NRMSE of J_{xy} is lower bounded to a significant value. In contrast, in conditions when J_{xy} is estimated with a better accuracy, the correct determination of J_{xy} is selected for all 100 runs of the procedure and the NRMSE of J_{xy} is no longer lower bounded: It regularly decreases when NK increases or when K decreases. This is precisely what occurs in the aforementioned example of Fig. 1 with $NK = 100\,000$: We manually checked all 100 estimates of J_{xy} (not shown here), which proved that one of them corresponds to a false determination (with a shift equal to a single step in the aforementioned grid) for $K = 200$ and no false determination for $K = 100$. The main conclusion of this analysis is that, when using enough state preparations, the proposed procedure avoids false determinations and, thus, has the usual behavior, with performance regularly increasing when the conditions (values of NK and/or K) are improved.

By considering a wide range of test conditions, Figs. 1 and 2 show that a wide range of estimation accuracies may be obtained for J_{xy} and J_z. Focusing on the most interesting cases, namely, when $K = 1$, the NRMSE of J_{xy} can, e.g., here be made equal to $2.75 \times 10^{-2} = 2.75\%$ for only $N = 10^4$ state preparations or 8.46×10^{-5} for $N = 10^5$ or 2.74×10^{-5} for $N = 10^6$. Similarly, when $K = 1$, the NRMSE of J_z can, e.g., be made equal to 7.66% for $N = 10^5$ or 2.17% for $N = 10^6$ or 9.07×10^{-5} for $N = 10^7$. The "very low" NRMSE values, corresponding to the absence of false determinations and to the parts "below possible discontinuities" in the plots of Figs. 1 and 2, are, thus, achieved for N higher than 10^4 for J_{xy} and 10^6 for J_z.

All above results show that, for given values of K and N, the parameter J_{xy} is often estimated much more accurately than J_z. This is reasonable for the following reason. J_{xy} is estimated by using only measurements along the Oz-axis, which lead to a relatively simple data model and, hence, a simple estimation procedure, which is likely to yield good estimation accuracy. In contrast, J_z is estimated by combining measurements along the Ox and Oz axes, and those along the Ox-axis involve a more complex data model, which yields an estimation procedure with possibly degraded estimation accuracy. It should also be noted that we here used a simple protocol by considering the same values of the set of parameters $\{K, N\}$ in the series of state preparations used for estimating J_{xy} and J_z. But one might instead use lower values of the number N of state preparations (preferably with $K = 1$) in the series of preparations performed for estimating J_{xy} than in those used for J_z, in order to balance the estimation accuracies achieved for J_{xy} and J_z while reducing the total number of state preparations (the BQPT method used here yields related considerations, which were detailed in [26]).

As explained in Section I, when considering parameter estimation tasks such as those related to BHPE, one may move from the most conventional configurations to the most advanced ones by considering three stages: One may first move from the classical to the quantum framework, then from nonblind to blind configurations, and finally from multiple-preparation methods to single-preparation ones. If aiming at comparing the performances of all these types of approaches, we already compared the last two, i.e., most advanced, approaches above: In the framework of blind quantum methods,

the results that we provided for $K = 1$ correspond to the most advanced, i.e., single-preparation, approach, whereas those for $K \neq 1$ (especially $K \gg 1$) correspond to the somewhat more standard, i.e., multiple-preparation, approach, although it still deals with methods that are advanced in the sense that they are blind, i.e., unsupervised.

Staying in the quantum framework, the remaining version of HPE that may be considered is the most conventional, i.e., nonblind, one. To perform such a comparison in a fair way, one should, however, keep in mind that blind methods anyway have the major qualitative advantage of not requiring known process input states, and that this might be obtained at the expense of lower accuracy. Having this difference in mind, a quantitative performance comparison may then be performed however. The question is then: Which nonblind HPE method should be used in this comparison? One might first think of using very generic nonblind HPE methods, i.e., methods that would apply to a large class of Hamiltonians, if any. However, using such a method in the present performance comparison might be considered not to be fair, because the price to pay for making such methods able to apply to a wide class of Hamiltonians, thus without requiring (nor using, if any) detailed prior knowledge about the considered Hamiltonian, is that these generic methods would most likely be very complex, as compared with the dedicated BHPE method that we proposed above for the Heisenberg model. Therefore, a more relevant comparison would be performed by testing a nonblind HPE method dedicated to the Heisenberg model, preferably with processing stages that are (nonblind but) similar to those used in our blind method, e.g., in terms of the types of spin components that are measured. To our knowledge, no such methods have been reported in the literature. However, in our paper [26] dealing with QPT instead of HPE, in addition to putting the emphasis on blind methods, we also introduced a nonblind QPT method dedicated to the process resulting from the Heisenberg Hamiltonian. We then compared the performance of our blind and nonblind QPT methods, thus showing that the blind one is more attractive. As for HPE methods, these results from [26] directly imply that our blind HPE method described above is more attractive than its nonblind counterpart, because both methods estimate the parameters J_{xy} and J_z that then define the process estimated by QPT methods, so that the performances of HPE and QPT methods are directly connected. For a more quantitative comparison, the reader is, therefore, referred to our results for QPT in [26].

B. CHANNEL ESTIMATION AND PHASE ESTIMATION

In Section I and Appendix A, we explained that, in the classical framework, the same information processing task is given different names, depending on the considered application field. In particular, the system identification task in the field of automatic control corresponds to the channel estimation task in the field of communications. The same phenomenon occurs in the quantum framework. In particular, QPT, and hence our blind (and possibly single-preparation) extension addressed in Appendix C, is often stated to be the quantum counterpart of classical system identification (see, e.g., [14, p. 389]). QPT applies to general quantum systems, not necessarily defined by a small set of parameters, and could, therefore, be called nonparametric system identification. But the HPE task, and hence our blind (and single-preparation) extension introduced in Section III-A, is also closely connected with system identification, and more precisely with parametric system identification. This results from the fact that it estimates a small set of parameters (e.g., the principal values of the exchange tensor in the case of Heisenberg coupling that was considered above as an example), and these parameters then completely define the behavior of that system, including the resulting process matrix in the associated QPT task.

Moreover, although a different terminology is used for other QIP tasks, some of these tasks actually address the same type of problems as above. This first concerns the quantum channel estimation task: As explained, e.g., in [63], a map from the density operator associated with a quantum state to another density operator is often called a quantum channel, as a reference to classical communication scenarios. The identification of such a map may, therefore, be called quantum channel estimation and is closely linked to the QPT problem that we considered above, possibly in its blind and single-preparation form. Similarly, a standard QIP procedure is phase estimation. In [14, p. 221], it is defined as the estimation of the phase Φ of an eigenvalue $e^{2\pi i\Phi}$ of a unitary operator. This task is, therefore, related as follows to both investigations reported in Appendix C and Section III-A. First, as explained in Appendix C, the considered (B)QPT problem essentially consists of estimating the parameters $\exp[i\frac{J_{xy}(t-t_0)}{\hbar}]$ and $\exp[i\frac{J_z(t-t_0)}{2\hbar}]$ and, hence, the exponential terms of the diagonal representation D of the considered operator [see (13)–(17)]. This is, therefore, equivalent to estimating the phases of these exponentials, up to a multiple of 2π. Moreover, the method introduced in Section III-A is directly connected with removing the additive indeterminacy due to this multiple of 2π.

This discussion shows that the blind and single-preparation extensions that we proposed above in this article for QIP tasks related to system identification are expected to be of importance not only for the scientific communities focused on QPT and HPE, but also for quantum scientists who investigate a variety of related problems, such as quantum channel estimation and phase estimation. Moreover, in this section, we restricted ourselves to problems related to the characterization (i.e., identification) of the considered quantum process itself. As explained in Section I and Appendix A, related QIP problems consist of building processing systems, with quantum and/or classical means, which essentially implement the *inverse* of an initially unknown quantum process. This corresponds to the quantum source separation (QSS) and related tasks, which we investigate in the next section, still aiming at *extending the considered configurations to blind and single-preparation ones.*

IV. QIP TASKS RELATED TO SYSTEM INVERSION AND STATE RESTORATION

A. BLIND QUANTUM SOURCE SEPARATION

A rather general version of the BQSS problem addressed here may be defined as follows. A set of qubits with indices j are independently prepared with states $|\psi_j\rangle$. The state $|\psi\rangle$ of the system composed of these qubits is, thus, equal to the tensor product of the above single-qubit states $|\psi_j\rangle$. This state $|\psi\rangle$ is then transformed, i.e., mapped to another state $|\psi'\rangle = \mathcal{M}(|\psi\rangle)$, where the mapping function $\mathcal{M}$, e.g., corresponds to temporal evolution with coupling between qubits, as detailed in the following. In the blind configuration, the user is given a set of transformed states $|\psi'\rangle$ but does not know the corresponding set of original states $|\psi\rangle$, and hence the source states $|\psi_j\rangle$, nor the mapping function $\mathcal{M}$. The user then eventually wants to restore the information contained in (at least part of) the source states. This information is obtained either in quantum form, by deriving estimates of these states $|\psi_j\rangle$, or in classical form, typically by eventually using a classical computer to derive estimates of the coefficients of the states $|\psi_j\rangle$ in a given basis.

This generic problem is connected with various application fields. The first one, on which we focus hereafter, is related to the operation of quamputers. In such a future quamputer, data will be stored in registers of qubits for subsequent use. Due to nonidealities of the physical implementation of such a register, the qubits that form it may have undesired coupling with one another, such as Heisenberg coupling, e.g., if considering quamputer implementations related to spintronics. As time goes on, the register state will, therefore, evolve in a complicated way due to this undesired qubit coupling, thus making the final value of that register state not directly usable in the target quantum algorithm executed on that quamputer. BQSS may then be used as a preprocessing stage, to restore the initial register state, before providing it to the target application of that quamputer.

To analyze this BQSS problem in more detail, we hereafter focus on a basic case, from which the reader may then extend this analysis to other configurations. In the considered case, the device (e.g., the qubit register) is restricted to two qubits, implemented as electron spins 1/2, and the undesired coupling that exists between them is again based on the cylindrical-symmetry Heisenberg model detailed in [26]. Using the notations of [26], the initial state $|\psi(t_0)\rangle$ of the device (e.g., the state stored at time t_0 in the register), which corresponds to state $|\psi\rangle$ in the above general definition of BQSS, may be represented by the column vector $C_+(t_0)$ of the components of $|\psi(t_0)\rangle$ in the standard basis, defined by

$$C_+(t_0) = [\alpha_1\alpha_2, \alpha_1\beta_2, \beta_1\alpha_2, \beta_1\beta_2]^T \tag{11}$$

where ^{T} stands for transpose. Similarly, the final state $|\psi(t)\rangle$ of the device (e.g., the only state available to the user, at a later time t, in the register), which corresponds to state $|\psi'\rangle$ in the above general definition of BQSS, may be represented by the column vector $C_+(t)$ of the components of $|\psi(t)\rangle$ in the standard basis. The effect of coupling is then represented by the relationship

$$C_+(t) = MC_+(t_0) \tag{12}$$

where M is a unitary matrix, which corresponds to the considered Heisenberg process and that may, therefore, be defined as follows [26], [41]:

$$M = QDQ^{-1} = QDQ \tag{13}$$

with

$$Q = Q^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{14}$$

and D equal to

$$\begin{bmatrix} e^{-i\omega_{1,1}(t-t_0)} & 0 & 0 & 0 \\ 0 & e^{-i\omega_{1,0}(t-t_0)} & 0 & 0 \\ 0 & 0 & e^{-i\omega_{0,0}(t-t_0)} & 0 \\ 0 & 0 & 0 & e^{-i\omega_{1,-1}(t-t_0)} \end{bmatrix}. \tag{15}$$

The four real (angular) frequencies $\omega_{1,1}$ to $\omega_{1,-1}$ in (15) depend on the physical setup. In [41], it was shown that they read

$$\omega_{1,1} = \frac{1}{\hbar}\left[GB - \frac{J_z}{2}\right]$$
$$\omega_{1,0} = \frac{1}{\hbar}\left[-J_{xy} + \frac{J_z}{2}\right] \tag{16}$$
$$\omega_{0,0} = \frac{1}{\hbar}\left[J_{xy} + \frac{J_z}{2}\right]$$
$$\omega_{1,-1} = \frac{1}{\hbar}\left[-GB - \frac{J_z}{2}\right]. \tag{17}$$

The first class of BQSS methods that we previously developed for handling this configuration (see especially [40]–[42]) is a quantum extension of classical ICA. This class is the "least quantum" one, in the sense that, starting from the available quantum states $|\psi(t)\rangle$, it first converts them into classical-form data (probability estimates) by means of measurements and then processes the latter data with only classical means. It, therefore, significantly differs from the new class of methods that we aim at introducing in this article since that new class mainly uses quantum processing means. We, therefore, skip the details of the aforementioned first class and refer the reader to, e.g., [40]–[42]. A major limitation of that class of methods should, however, be mentioned: It requires each state $|\psi(t_0)\rangle$ to be prepared many times, both in the so-called adaptation and inversion phases, whose definitions are detailed further for the new class of methods proposed in this article and may be summarized as follows (this also corresponds to the general features that we provide in Appendix A for classical and quantum machine learning methods). The adaptation (or training) phase is first used

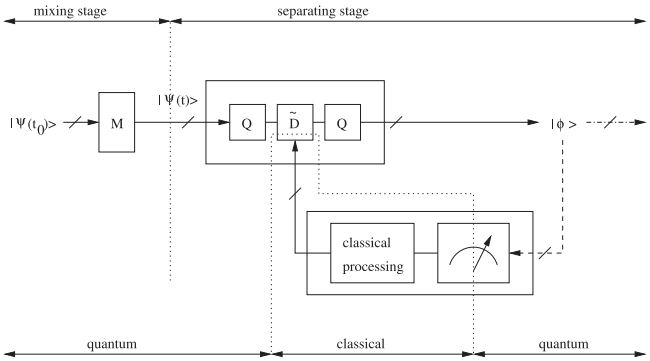

FIG. 3. Global (i.e., mixing + separating) configuration, with a feedback separating system that includes a quantum-processing inverting block and a classical-processing adapting block. Each quantum state $|\Phi\rangle$ is used only once (no cloning): See [45].

to select an adequate behavior for the proposed separating system, i.e., to learn the (direct and) inverse mapping. The inversion phase then corresponds to the final, useful, operation of that system, after learning. The aforementioned need for many copies of the states $|\psi(t_0)\rangle$ makes our first class of BQSS methods "less blind" because, although these states $|\psi(t_0)\rangle$ are allowed to be unknown from the point of view of the adaptation procedure, some control is required so that the *same* value is repeatedly prepared for each of these states.

A solution to the above problem was introduced, but only for the inversion phase, in our second class of BQSS methods, especially described in [43]–[45]. We now detail it since we take advantage of it in the new, fully SIPQIP, methods that we introduce further in this article for BQSS. The complete operation of that second class of BQSS methods again consists of the aforementioned two successive phases. We here first describe the inversion phase. During that phase, each state $|\psi(t)\rangle$ available as the input of the separating system is directly used in quantum form, i.e., without performing measurements, so that this separating system outputs a quantum state $|\Phi\rangle$ that should ideally be equal to the multiqubit source state $|\psi(t_0)\rangle$ that one aims at restoring. That part of the separating system, called the inverting block, is thus a global quantum gate, which corresponds to the upper right box in Fig. 3 (that is, it consists of the cascade of (sub)gates Q, $\tilde{D}$, and Q, as justified hereafter). This inverting block, thus, only requires a *single* instance of its input state $|\psi(t)\rangle$ to derive its corresponding output state $|\Phi\rangle$.

This inverting block is designed as follows. We exploit the fact that, although the actual value of the mixing matrix M of (12) is not known in the blind configuration, one knows that it belongs to the class of unitary matrices defined by (13), where $Q = Q^{-1}$ is a known, fixed, matrix and D is a diagonal matrix, whose diagonal entries have unit modulus (and a structure that is disregarded in this approach), as shown by (14) and (15). We, therefore, use an inverting block of the separating system, which is adaptive (or tunable), i.e., such that some of the values of the parameters that define its behavior may be modified. More precisely, this block is designed so that it is able to implement the inverse of any transform in the above-defined class, depending on its parameter values. Its operation is, therefore, represented by a unitary matrix defined as

$$U = Q\tilde{D}Q \tag{18}$$

with

$$\tilde{D} = \begin{bmatrix} e^{i\gamma_1} & 0 & 0 & 0 \\ 0 & e^{i\gamma_2} & 0 & 0 \\ 0 & 0 & e^{i\gamma_3} & 0 \\ 0 & 0 & 0 & e^{i\gamma_4} \end{bmatrix} \tag{19}$$

where γ_1 to γ_4 are free real-valued parameters. This inverting block is, thus, the cascade of three simpler quantum gates, as shown in the upper right box of Fig. 3. The implementation of each gate corresponding to the matrix Q, as a combination of even simpler gates, was detailed in [41]. Moreover, the adaptive gate corresponding to (19) was introduced in [45], and its implementation was also detailed there. In [45] and in the new use of that gate introduced further in this article, the values of the parameters γ_1 to γ_4 are controlled by classical-form signals. These parameters may, e.g., be independent, known but arbitrary, increasing, functions of control voltages. Such control voltages are, e.g., used in the real device described in [64].

Before the above-defined inversion phase, the complete operation of that second class of BQSS methods, therefore, requires an adaptation phase, during which a set of states $|\psi(t)\rangle$ is used to adapt the matrix $\tilde{D}$, i.e., to learn the inverse mapping. The method used to this end in [45] is based on the probabilities of measurements associated with a set of output states $|\Phi\rangle$ of the inverting block of Fig. 3. These probabilities are essentially used to measure the degree of entanglement of these states $|\Phi\rangle$. The matrix $\tilde{D}$ is adapted so as to essentially make these states $|\Phi\rangle$ unentangled so that this type of methods performs an "unentangled component analysis" [45], as opposed to the aforementioned classical and quantum PCA and ICA. The complete structure of the resulting separating system is shown in Fig. 3. Unlike in the inversion phase, during the adaptation phase, this structure requires many copies of each of its input states $|\psi(t)\rangle$, in order to derive the corresponding copies of the output states $|\Phi\rangle$ and, hence, the corresponding probability estimates based on sample frequencies of measurement outcomes. This second class of BQSS methods is, thus, "more quantum" than the first one, first because it uses quantum processing means in the inverting block, and second because it is based on the quantum concept of entanglement, which has no classical counterpart.

In this article, we introduce a third and new class of BQSS methods. This new class proceeds further than the previous two, by using the SIPQIP framework in all the operation of the separating system, i.e., by using a single preparation of each state also during the adaptation phase. To this end, we exploit the structure of the matrix D, defined in (15)–(17). We

IEEE Transactions on Quantum Engineering

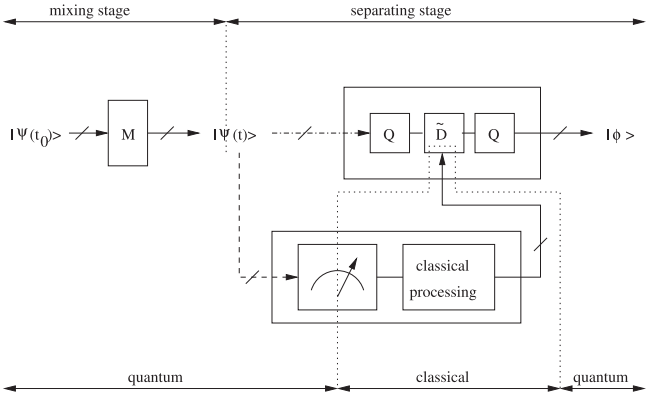

FIG. 4. Global (i.e., mixing + separating) configuration, with a feedforward separating system that includes a quantum-processing inverting block and a classical-processing adapting block. Each quantum state $|\psi(t)\rangle$ is used only once (no cloning): See text.

take into account the fact that the matrix $\tilde{D}$ of the separating system should ideally be set to the inverse of D. Therefore, by replacing J_{xy} and J_z by their estimates $\widehat{J}_{xy}$ and $\widehat{J}_z$ in (15)–(17), we set $\tilde{D}$ as in (19), but here with the following structure for the phases of its diagonal elements:

$$\gamma_1 = \frac{GB\tau_3}{\hbar} - \frac{\widehat{J}_z\tau_3}{2\hbar}$$
$$\gamma_2 = -\frac{\widehat{J}_{xy}\tau_3}{\hbar} + \frac{\widehat{J}_z\tau_3}{2\hbar} \tag{20}$$
$$\gamma_3 = \frac{\widehat{J}_{xy}\tau_3}{\hbar} + \frac{\widehat{J}_z\tau_3}{2\hbar}$$
$$\gamma_4 = -\frac{GB\tau_3}{\hbar} - \frac{\widehat{J}_z\tau_3}{2\hbar} \tag{21}$$

where τ_3 is the value of the time interval $(t - t_0)$ used in the inversion phase of the proposed BQSS method. In this new BQSS method, we take advantage of the BQPT method that we described in Appendix C. When applying the latter method with time intervals $(t - t_0)$ set according to (29) with a freely selected value of τ_1, we get (44) and (46), which shows that all the quantities in (20) and (21) required to assign γ_1 to γ_4 are known or can be estimated. The adaptation phase of the proposed BQSS method, therefore, consists of the following three steps. One first applies the above BQPT method. One then uses (20) and (21) to derive the selected values of γ_1 to γ_4. One finally uses the supposedly known correspondence function, which makes it possible to convert these values of γ_1 to γ_4 into the practical control signals (e.g., voltages) of the gates implementing the matrix $\tilde{D}$, which make these gates operate with these desired values of γ_1 to γ_4. The resulting global configuration is shown in Fig. 4. Each state $|\psi(t)\rangle$ is, thus, used only once (see also [45] about the no-cloning theorem). During the adaptation phase, these states are sent to the part of the system, which performs measurements and BQPT (dashed line and lower part of Fig. 4); then, during the inversion phase, they are sent to the inverting block (dash-dotted line and upper right part of Fig. 4).

Thanks to the properties of the BQPT method reused here as a building block of the proposed BQSS method, the output of the latter method does not depend on the values used in (44) and (46) for the integers $\widehat{k}_{xy}$ and $\widehat{k}_z$. This may be seen by inserting (44)–(46) in (20) and (21) and then in (19), with (29), which shows that the terms of (44)–(46) that include $\widehat{k}_{xy}$ and $\widehat{k}_z$ yield terms in (20) and (21) that are integer multiples of 2π, and that, therefore, have no influence on the value of $\tilde{D}$.

B. BLIND QUANTUM (ENTANGLED) STATE RESTORATION

In the classical framework, the concept of BSS intrinsically refers to situations involving several (unknown) source signals, created by several "sources" that may be some kinds of "objects." Such a situation may be mathematically described by gathering all the values of these source signals, e.g., at a given time, as the elements of an overall source vector. In our quantum extensions of this classical BSS, we started from a similar situation, involving several objects, such as qubits implemented as spins 1/2, and we first independently considered the "signal value," i.e., the initial quantum state $|\psi_j(t_0)\rangle$, of each of them [see (30)]. We then "gathered" these individual states by defining the state of the complete system [see (31)] as the tensor product of the states $|\psi_j(t_0)\rangle$, somehow as the quantum counterpart of the above vector of classical signal values. However, the quantum framework opens the way to much richer situations. This is due to the fact that the possible states of a complete system are not restricted to the tensor products of the individual states of independent parts of this system: They also include entangled states. An extension of the above BQSS problem is, therefore, blind quantum state restoration (BQSR), aiming at restoring a *possibly entangled* deterministic-coefficient pure state of a multiqubit system, starting from an altered version of it. One, thus, conceptually considers a single arbitrary multiqubit source state, instead of several single-qubit source states. In particular, this includes restoring the possibly entangled initial state $|\psi(t_0)\rangle$ of a multiqubit system, from the state $|\psi(t)\rangle$ of that system at a later time.

We here investigate this extension of BQSS to possibly entangled source states, by considering the use of such states in the second phase of the operation of this system, i.e., in the above-defined inversion phase, after the transform performed by this system has been fixed by means of the adaptation phase. We claim that the BQSS system defined in Fig. 4 is directly able to perform the considered BQSR task, since it operates as follows. An unknown state $|\psi(t_0)\rangle$ is created and then modified by an operator represented by the matrix M, thus yielding the state $|\psi(t)\rangle$. The latter state is the state processed by the new separating system that we designed in Section IV-A. The transform performed by this system is represented by the matrix $U = Q\tilde{D}Q$ [see (18)]. But, during the adaptation phase of that separating system that occurred before this inversion phase, U was made equal to the inverse

of M (up to estimation errors) by the proposed adaptation method. Therefore, when applying that separating transform U to $|\psi(t)\rangle$, the output state of the separating system becomes equal to $|\psi(t_0)\rangle$ (up to estimation errors), and this analysis does not depend at all whether $|\psi(t_0)\rangle$ is entangled or not. In particular, if the proposed BQSS (and hence BQSR) method is used to restore the initial state of a qubit register by compensating for the undesired coupling between its qubits, as discussed above, this means that this method also applies when an entangled state is stored in this qubit register.

C. BLIND QUANTUM CHANNEL EQUALIZATION

As discussed in Sections I and III-B and in Appendix A, in the classical and quantum frameworks, the same information processing task is given different names depending on the considered application field. This is also true for the generic QIP problem, related to system inversion, that we initially defined for nonentangled states at the beginning of Section IV-A, that we then extended to possibly entangled states in Section IV-B, and that may be summarized as follows: A user is given a set of transformed states $|\psi'\rangle = \mathcal{M}(|\psi\rangle)$, but does not know the original states $|\psi\rangle$, nor the mapping function $\mathcal{M}$; the user wants to restore the original states $|\psi\rangle$. Whereas we illustrated that QIP task with one type of application in Sections IV-A and IV-B, we anticipate that it will have various other applications as the quite general field of QIP keeps on growing. In particular, this problem too may be rephrased as a quantum communication scenario. The above user is then the receiver, who only knows the received states $|\psi'\rangle$, which have been altered by the channel $\mathcal{M}$. Without knowing that channel, the receiver aims at restoring the emitted states $|\psi\rangle$ that he does not know either (or the information, in classical form, contained in these states $|\psi\rangle$). This, therefore, corresponds to the blind quantum channel equalization problem. The generic methods that we proposed in Sections IV-A and B also allow one to solve this problem, moreover with the advantages of the proposed SIPQIP framework.

V. CONTRIBUTING TO QUANTUM CLASSIFICATION

We now come back to the other main aspect of machine learning discussed in Appendix A, namely, classification. In the classical framework, many classification algorithms receive data that consist of vectors, which contain features that characterize the "objects" to be classified [65]. These algorithms heavily rely on computing the dot (i.e., scalar or inner) product $v_j^T v_k$ of two column vectors v_j and v_k, or on computing the distance $||v_j - v_k||$ between the associated two data points [65]. These two quantities are, moreover, directly connected, since

$$||v_j - v_k||^2 = ||v_j||^2 + ||v_k||^2 - 2v_j^T v_k. \tag{22}$$

In particular, for any unit-norm vectors, this yields

$$||v_j - v_k||^2 = 2(1 - v_j^T v_k). \tag{23}$$

Using the signal/data processing terminology, $v_j^T v_k$ is also the basic, i.e., noncentered and nonnormalized, correlation parameter of the data vectors v_j and v_k, whereas their noncentered correlation coefficient (also called the cosine similarity [65]) is

$$\rho(v_j, v_k) = \frac{v_j^T v_k}{||v_j||.||v_k||}. \tag{24}$$

These two correlation parameters coincide for unit-norm vectors. The correlation coefficient $\rho(v_j, v_k)$ is, e.g., widely used for data characterization and classification by the Earth observation (i.e., remote sensing) community. Each of the vectors v_j and v_k then typically defines a spectrum, which consists of the light reflectance values of a material at a set of "frequencies" (in fact, narrow spectral bands) [34]. More precisely, one then computes the arccosine of $\rho(v_j, v_k)$, which is equal to the angle between v_j and v_k [65] and is, therefore, called the "spectral angle mapper" (SAM) between these vectors [66]. A low value of that SAM corresponds to a high value of the "spectral similarity" of the considered materials, i.e., of their similarity in terms of the shape of the variations of their reflectance functions with respect to frequency (where this shape does not refer to their global scale: That scale has no influence on (24) and, hence, on SAM). Similar approaches are used in the field of Astrophysics, with spectral data vectors that consist of luminance values (i.e., direct light flux from the observed object), instead of reflectance.

Let us now consider the situation when data that are initially in classical form are to be classified by using a quantum classifier, in order to achieve higher classification speed [10], [65], [67]. This first requires one to transform the initial classical data into quantum states. To this end, each classical-form vector is stored in a qubit register with index r, which consists of Q qubits. Each individual qubit is, thus, indexed by r and $q \in \{1, \ldots, Q\}$. Its state space is denoted as $\mathcal{E}_{rq}$, and a basis of this space is composed of the two kets $|k_{rq}\rangle_{rq}$ with $k_{rq} \in \{0, 1\}$. All deterministic-coefficient pure states of the qubit register r then belong to the space $\mathcal{E}_{r1} \otimes \ldots \mathcal{E}_{rQ}$ and read

$$|\psi_r\rangle = \sum_{\mathcal{S}(k_{r\bullet})} c_{rk_{r1}\ldots k_{rQ}} |k_{r1}\rangle_{r1} \otimes \cdots \otimes |k_{rQ}\rangle_{rQ} \tag{25}$$

where $\otimes$ is the tensor product and the compact notation $\mathcal{S}(k_{r\bullet})$ means the set of all values of the ordered set of Q integers k_{rq} corresponding to the fixed value r and to all values $q \in \{1, \ldots, Q\}$, again with $k_{rq} \in \{0, 1\}$. Besides, the 2^Q complex-valued coefficients $c_{rk_{r1}\ldots k_{rQ}}$ are indexed by the index r of the considered register and by all integers k_{rq}, which define to which basis state each coefficient $c_{rk_{r1}\ldots k_{rQ}}$ corresponds. These coefficients are such that $|\psi_r\rangle$ has unit norm. Let us then consider a classical-form complex-valued unit-norm vector v_j with dimension 2^Q (or lower: zero-valued components are then added to v_j to reach 2^Q components). This vector may be stored in a ket $|\psi_r\rangle$ defined by (25), by setting the coefficients $c_{rk_{r1}\ldots k_{rQ}}$ of $|\psi_r\rangle$, respectively, to the values of the components of v_j (a common phase reference may be used for all considered kets). If the norm of v_j is not

equal to one, it may be handled separately, while $v_j/||v_j||$ is stored in $|\psi_r\rangle$, as stated in [67].

The above-defined kets (25) may then be employed in quantum classifiers, which often use: 1) the dot product $\langle\psi_1|\psi_2\rangle$ of such kets; 2) the squared modulus of this dot product, which is called the overlap of these kets (see [65, p. 120], [68], or [69]); or 3) the distance between the points associated with such kets, as, e.g., discussed in [10], [65], and [67]. The dot product formally associated with two states (25), respectively, stored in registers with indices $r = 1$ and $r = 2$ is defined as if these kets belonged to the same state space. This dot product, therefore, reads

$$\langle\psi_1|\psi_2\rangle = \sum_{\mathcal{S}(k_{r\bullet})} c^*_{1k_{r1}\ldots k_{rQ}} c_{2k_{r1}\ldots k_{rQ}} \tag{26}$$

where * stands for complex conjugate. Some quantum circuits were proposed in the literature for computing the corresponding state overlap. A widely used approach, called the swap test, was proposed in [70] to essentially test the equality of two states. The quantity used to this end is the probability of an outcome of a measurement performed at the output of the considered circuit. This quantity is equal to 0 if the considered states are equal and essentially equal to 1/2 otherwise (more precisely, it is higher than a bound close to 1/2 if the states are far enough from one another). Beyond this binary decision, this probability is a continuous-valued quantity, which may be shown to be linearly related to the overlap of the considered quantum states (part of the corresponding calculations are provided in [65] and [70]). Another approach for computing the overlap of quantum states is based on the circuit of [68, Fig. 6(B)]. The behavior of that circuit is only briefly defined in [68]. That article outlines how to express the overlap associated with the density operators of the two multiqubit inputs of the considered circuit as the result of classical postprocessing applied to the results of measurements performed at the output of that circuit. Our own calculations (to be detailed elsewhere), performed for deterministic-coefficient pure input states $|\psi_1\rangle$ and $|\psi_2\rangle$, confirm that the squared modulus of (26) may be expressed as a (multistage) linear combination of probabilities of outcomes of measurements performed at the output of that quantum circuit. In Appendix F, we show how the above quantum circuits may be further exploited in order to compute dot products $\langle\psi_1|\psi_2\rangle$, not only their (squared) moduli.

The above dot products or overlaps may be used in various ways in the general framework of quantum classification. More specifically, we, hereafter, show how enhanced approaches may be developed by combining quantum classification principles that use dot products or overlaps with our SIPQIP, i.e., single-preparation, concept. We illustrate this approach with a first original contribution to single-preparation quantum classification, which will be extended in future papers. In this contribution, we focus on the second phase of the operation of a classifier, that is on the "resolution phase" defined in Appendix A, which takes place after the (unsupervised or supervised) learning phase. This is, by the way, similar to what we did for BQSS, by first applying our SIPQIP concept to the second phase of operation, i.e., the inversion phase, before we extended it to the first, i.e., adaptation, phase, as explained in Section IV-A.

In the proposed approach, we consider the situation when the classical-form vectors to be classified are characterized by their shapes, not their magnitudes, e.g., as in the Earth observation and Astrophysics applications outlined at the beginning of the present section. Therefore, these vectors may initially be rescaled to have unit norm, so that this norm is not an issue when transforming these classical vectors into quantum states. We, hereafter, address the general situation when the considered classification problem involves C classes, indexed by c, with $c \in \{1, \ldots, C\}$. Moreover, we consider the usual case when the classical-form data vectors, and, hence, the associated dot products of kets, are real-valued.

To describe how classification is here performed, let us first consider the nonrealistic situation when each class with index c is initially defined by a single known classical-form vector v_{c1} and, hence, a single associated quantum state $|\psi_{c1}\rangle$. We consider the analysis of a new "object" of the considered application (e.g., the spectrum of an unknown material in the above Earth observation or Astrophysics applications), represented by a quantum state $|\phi\rangle$. A basic method for classifying that object consists of separately estimating its dot product $\langle\phi|\psi_{c1}\rangle$ with each of the states $|\psi_{c1}\rangle$ and in deciding that this object belongs to the class that yields the highest estimated value of the dot product $\langle\phi|\psi_{c1}\rangle$, i.e., the best similarity with $|\phi\rangle$. This approach may be simplified as follows in the case when the components of the considered classical-form data vectors are nonnegative, which, e.g., applies to the reflectance or luminance values that compose the aforementioned spectra. In that case, the square root of the overlap $|\langle\phi|\psi_{c1}\rangle|^2$ coincides with the corresponding dot product $\langle\phi|\psi_{c1}\rangle$. This overlap is, therefore, sufficient for measuring similarity in that case (as opposed to the sign indeterminacy that it yields with respect to the dot product for possibly negative data). The above classifier then operates equivalently by deciding that the considered object belongs to the class that yields the highest estimated value of overlap $|\langle\phi|\psi_{c1}\rangle|^2$. This is attractive because an overlap is computed more easily that the corresponding dot product, as shown in Appendix F.

An improved variant of the above classification method employs a user-defined threshold in addition, in order to achieve the rejection capability defined in Appendix A. In this approach, if the highest of the above dot products (or overlaps, in the simplified version) remains lower than this threshold, the considered object is "rejected." This means that the classifier then decides that it is not able to classify that object because it is not similar enough to any of the classes of objects that are known in the considered problem.

All these classifiers are based on computing overlaps, because their decisions are either directly based on such overlaps or based on dot products, which may be derived from

overlaps, as explained in Appendix F. Each of these overlaps, such as $|\langle\phi|\psi_{c1}\rangle|^2$, is typically estimated by using the sample frequency estimate(s) of one or several types of probabilities associated with overlap in the quantum circuits that were defined above for estimating overlaps. This then means that, for each class c, many copies (typically 10^5, as explained in Appendix B) of the state $|\psi_{c1}\rangle$ must be prepared to estimate these probabilities.

Now, consider the realistic version of the above problem, when each class with index c is initially defined by a full set of classical-form vectors v_{cj}, with a vector index j ranging from 1 to a maximum value that may depend on the class. These vectors are then transformed into quantum states $|\psi_{cj}\rangle$. The above classifiers may be extended as follows for this situation, focusing on their version that directly bases its decisions on overlaps, for the sake of clarity. For each class, one may first compute a full set of (estimates of) overlaps $|\langle\phi|\psi_{cj}\rangle|^2$ and these quantities should then be reduced to a single parameter that characterizes the overall similarity of the considered class with $|\phi\rangle$. A natural parameter that may be used to this end is the mean of all overlaps $|\langle\phi|\psi_{cj}\rangle|^2$ associated with the considered class. Here again, in practice, only an estimate of this mean overlap is obtained, by using various quantum state preparations and measurements. However, unlike in the above nonrealistic scenario, this may here be achieved with two quite different approaches. The first approach, which might be considered as the most natural one if disregarding our previous contributions in this article, consists of separately estimating each of the overlaps $|\langle\phi|\psi_{cj}\rangle|^2$ as above, therefore typically preparing 10^5 copies of each state $|\psi_{cj}\rangle$ (and of $|\phi\rangle$), and then computing (on a classical computer) the mean of these estimated overlaps. This is an application of the standard, multiple-preparation, approach defined in Appendix B. However, we stress that we here only aim at computing the *mean* of this finite set of overlaps so that we only need to estimate the mean(s) of the corresponding set(s) of probabilities (as explained above, this involves one or several types of probabilities, depending on the considered quantum circuit). In Section II-B, we showed that this may be performed much more efficiently by using our SIPQIP framework. This here means decreasing the number of preparations per state $|\psi_{cj}\rangle$ and taking advantage of the averaging of measurement results that is then performed over all these states (thus still requesting one copy of $|\phi\rangle$ per measurement). This number of preparations per state may even be decreased down to one if enough different states $|\psi_{cj}\rangle$ are available to reach a high enough estimation accuracy: In [26] and Section III-A2 of the present article, we analyzed the numerical performance achieved by this SIPQIP approach for the BQPT and BHPE tasks, and we plan to investigate it for classification in future papers. In the literature, quantum classifiers have especially been considered for big-data, i.e., large-scale, applications [10], [67]. In such applications, the aforementioned large number of states $|\psi_{cj}\rangle$ will actually be available and our SIPQIP framework will take full advantage of it (besides, it can also attractively operate with a somewhat lower total number of states $|\psi_{cj}\rangle$ and a number of preparations per state somewhat higher than one).

VI. CONCLUSION

The term "machine learning" especially refers to algorithms (and associated systems) that derive mappings, i.e., input–output transforms, by using numerical data that provide information about the transform, which is of interest in the considered application. The data processing tasks to be performed in these applications not only include classification and regression but also system identification, system inversion, and input signal restoration (or source separation when considering several signals). Whereas these problems have been and are still widely investigated in a purely classical framework, part of them are currently being extended to configurations, which involve quantum-form data and/or quantum processing means. Within this general quantum framework, we here tackled the most challenging configurations from two points of view. First, almost all this article is devoted to *unsupervised*, i.e., blind, configurations, which have not been addressed in the literature for most of the tasks considered here. Unsupervised learning is very attractive because, as detailed in Section I and Appendix A, it avoids the need for known "reference values" (e.g., input values for system identification) to learn the required mappings. Second, we here mainly aim at extending a variety of aspects of quantum machine learning by introducing new algorithms, which can operate with only one instance of each prepared state (where the term "preparation" is used for both deterministic-coefficient and random-coefficient pure states, as explained in Section I). This approach first avoids the burden of having to prepare many ideally identical copies of each used state in order to compute statistical parameters separately for each such state. Moreover, this approach yields much better performance than the multiple-preparation approach for a given total number of state preparations, as shown for BQPT in our very recent paper [26] and confirmed here by our new results for BHPE. Our original single-preparation approach also compares favorably with the multiple-preparation approach from the following point of view, in the framework of unsupervised learning. Using the multiple-preparation approach in that framework means allowing the "reference values" to be unknown but still requesting that the *same* (unknown) reference value be prepared many times. This still requires significant control in the considered quantum learning procedure so that this procedure is "less unsupervised." Our single-preparation approach avoids this problem and is, therefore, attractive.

The above concepts, thus, result in a general SIPQIP framework. Whereas our recent papers [26], [50] introduced that single-preparation concept itself but only showed how to apply it to a single task (BQPT), our new contributions in this article consist of describing how it can be attractively used to perform a large range of other QIP tasks: BHPE, blind quantum channel identification/estimation, blind phase estimation, BQSS, blind quantum entangled state restoration,

and blind quantum channel equalization, classification. We, moreover, provide a quantitative evaluation of the resulting numerical performance. For the tasks related to the blind, i.e., unsupervised, version of system identification (including HPE), system inversion, and signal restoration (including source separation), we applied the proposed approach to a concrete example, related to spintronics, which involves Heisenberg coupling between two qubits. Starting from the explicit algorithms and system architectures that we detailed for this configuration, the reader may then adapt them to other types of processes. We will also extend this approach to other processes in the future. Similarly, we provided a first illustration of the application of this SIPQIP framework to quantum classification, and we plan to report extensions of this approach in future papers.

Moreover, when aiming at compensating for undesired Heisenberg coupling between qubits, we defined two quantum system architectures: See Fig. 3 for a feedback structure and Fig. 4 for a feedforward structure. These architectures open the way to the much more general concept of "self-adaptive quantum gates," i.e., gates with two features:

1) some means for controlling the values of parameters that define the quantum state transform that such a gate performs within a predefined class of transforms;
2) an autonomous (i.e., blind or unsupervised) algorithm, which controls the adaptation of these parameter values, so as to achieve a predefined condition. That condition could consist of ensuring output disentanglement, as in the above example. It could instead be a counterpart of disentanglement, depending on the available data and on the type of undesired behavior that one wants to compensate for.

Such gates would especially be of interest for a *qua*ntum c*omputer* that, by the way, we proposed to more briefly call a "quamputer": By adequately selecting the aforementioned adaptation condition and designing an associated adaptation algorithm, one could create a self-adaptive quantum gate that automatically compensates for a given type of nonideality (instead of undesired Heisenberg coupling in the above example) that occurs, e.g., in a gate that precedes the considered self-adaptive gate. This would allow practical future quamputers to operate correctly despite these nonidealities, thanks to their internal compensation means.

APPENDIX A
CLASSICAL AND QUANTUM MACHINE LEARNING APPROACHES FOR DATA MAPPING

Many classical and quantum information processing systems aim at applying transforms, defined by mathematical functions, to their input data, in order to map them to output data. These transforms are often called mappings or maps, both in the classical [6] and quantum [63] information processing literature (quantum maps are also called quantum channels [63], with a reference to communications). In basic systems, the considered transform is predefined by the human system designer, depending on the target application. In contrast, more advanced systems, which are considered hereafter, are referred to as (self-)adaptive systems since they adapt their behavior (i.e., the mapping they perform) to the data they receive [71], [72], by means of algorithms that perform the so-called adaptation, training, or (machine) learning [1], [6]–[9]. In other applications, input–output mappings are also learnt from data but with other goals, especially to characterize the behavior of a given natural medium or artificial system, as detailed in the following. The classical and quantum versions of machine learning, thus, involve various types of applications and associated types of transforms, which are analyzed in more detail hereafter.

Machine learning is first used in classical classification and regression systems, whose transforms map a set of input quantities (each of which has its own nature) to output quantities, which often have a different nature from input ones. This is especially true for classification systems [1], [6]–[9], which receive a set of input quantities that are most often continuous valued, whereas their (possibly thresholded) outputs are binary valued. More precisely, let us first consider a classification system without the rejection capability that is defined further. Such a system generally outputs C values, where C is the number of classes involved in the considered application. Only one of these outputs is equal to 1, say output with index c, whereas all other outputs are equal to 0. The index c of the active output defines the decision made by the classifier: It considers that its input values correspond to a case when the input belongs to class c. During the final use of the classifier, called the "resolution phase," "classification phase," or "test phase," the above output values are provided to the target application. A typical use of this framework is optical character recognition (OCR) [6], [8], [9]. The classifier then receives an image, i.e., a set of pixel values (or features, i.e., parameters, extracted from them), where a letter or symbol belonging to a given alphabet is written. The classifier sets its cth output to 1 if it considers that this particular input image contains the cth symbol of that alphabet. Moreover, improved classification algorithms are able to detect when they consider that the input that they receive during the resolution phase does not belong to any of the considered classes, e.g., when the received image is not similar enough to any character of the considered alphabet (indeed, an image may contain a shape that is not a character of any written language). Such a classifier then decides that it is not able to classify the considered input "object" and it rejects it. This may be expressed, e.g., by setting all C outputs to zero or by adding a $(C+1)$th output to the classifier, which is equal to 1 when this classifier succeeds in classifying the considered input and to 0 otherwise.

Before the above resolution phase, machine learning algorithms are typically used, e.g., in OCR systems, to initially build an adequate input–output mapping, during the so-called learning phase, training phase, or adaptation phase (this possibly includes a so-called validation), by using data

composed of training examples. In supervised learning approaches, each example consists of an input (e.g., an image containing a character for OCR) and its correct class, i.e., the associated desired values of the classifier outputs (called labels), which are provided by a supervisor, i.e., typically a human expert of the considered application. In contrast, in unsupervised learning approaches for classification (called clustering [9]), the system self-organizes by using only inputs (e.g., images in the OCR example), i.e., unlabeled data. Various system architectures and supervised or unsupervised learning algorithms have, thus, been developed, especially including (artificial) neural networks [6]–[9] and their recent deep extensions [1], as well as support vector machines (SVMs) [7]–[9].

Regression systems [6], [7], [9] are similar to the above classifiers, except that their outputs are continuous-valued. They typically first use a supervised training phase in order to learn mappings from data samples, which are examples of adequate pairs composed of input values and correct corresponding output values in the considered applications. Once this mapping has been fixed, such a regression system may eventually be used, e.g., to control an industrial setup in a factory: The regression system then receives, as its inputs, different types of measured quantities provided by sensors available in the factory, and this system maps its inputs to possibly different types of continuous-valued quantities, used to drive the actuators that control the industrial setup. More specifically, regression systems where the inputs and output(s) have the same nature especially concern prediction tasks for time series, where the system aims at providing the expected future value(s) of a quantity (e.g., a currency exchange rate or streamflow) from its past values.

Whereas the above concepts were initially developed for classical data, they are currently being extended to quantum data and/or quantum processing means [3], especially because the QIP [14] community is investigating quantum extensions of classifiers to handle the huge processing power and amount of data involved in current real-world applications. These extensions include the implementation of SVM classifiers on quantum computers with very low computational complexity [10]. Besides, the versatile quantum optical neural networks proposed in [73] can perform different related tasks, including reinforcement learning.

Although one may first have in mind the above general classification and regression/prediction tasks when thinking of classical and quantum machine learning, data-driven algorithms are also widely used in a partly related set of processing tasks, called system identification and system inversion, with an extension to source separation. First considering the classical framework, this, e.g., includes situations when an electromagnetic or acoustic signal is emitted from a first location, then transferred through a medium, which may be seen as an electromagnetic or acoustic "channel" that transforms its input composed of the emitted signal. The output of that channel is then the signal received by an electromagnetic antenna or microphone in a second location. These and other practical situations yield two types of machine learning problems. The first one is often called system identification [11]–[13] and, more specifically, channel identification or channel estimation in the field of electromagnetic communications [12]. Its simpler version [13], called the "nonblind version" by the signal and image processing community and the "supervised version" by the machine learning and data analysis community, operates as follows. As in the above regression task, it uses a set of known continuous-valued pairs composed of input values and corresponding output values of the considered "system," i.e., of the above-defined channel, in order to estimate, i.e., learn, the unknown transform (i.e., mapping) performed by this system. It should be noted that, in the above examples involving electromagnetic or acoustic signals, the considered "system" is not an artificial system to be built by human beings in order to perform a given type of data processing (as in the above classification and regression tasks), but the "natural system" formed by the considered electromagnetic or acoustic propagation medium. The goal of system identification is then to characterize this medium.

System identification methods have also been extended to the more challenging blind, i.e., unsupervised, configuration [11], [12]. In that case, during the learning or estimation phase, only the values of the system output (i.e., the values of the received channel output in the above examples) are known, whereas its input values are unknown. However, the system input is most often required to have some known properties, e.g., some known statistical features so that this configuration is sometimes stated to be semiblind or semisupervised. This problem may, therefore, be seen as a nonconventional form of regression, with only partial knowledge about the input data. Besides, it should be noted that the known values are only the *output* values in blind system identification, whereas they are only the *input* values in the above unsupervised classification problem. Whereas nonblind or blind system identification is applied to single-input single-output systems in its basic form, it may then be extended to multiple-input multiple-output (MIMO) systems.

The second considered machine learning problem, in connection with system identification, deals with system inversion and signal restoration. One then again considers an unknown "system," called the direct system, but one here aims at building an artificial system that essentially performs a transform equal to (an estimate of) the inverse of that of the direct system (assuming that direct transform is invertible). This inverse transform is first learnt during the training/adaptation phase, either directly or by first learning the direct transform (using system identification methods) and then deriving its inverse (in low-noise scenarios). Then, in the "inversion phase," which corresponds to the final use of the inversion system, the (estimated) inverse transform is applied to known output values of the direct system, in order to recover (estimates of) corresponding unknown input values of that direct system. This approach, e.g., applies to the

above two configurations involving channels. The direct system then corresponds to the physical propagation medium, such as an electromagnetic communication channel, which alters the emitted signal in a initially unknown way. One then aims at restoring the emitted signal from its altered received version, by learning an adequate inverse transform from data samples. In the field of radio frequency communications, this signal processing task is often referred to as (channel) equalization [12], [46]. Similarly, the restoration of an unknown emitted acoustic signal from its received version altered by reverberation during propagation is often called dereverbation [47]. More generally speaking, the problem of restoring a source signal only from an observation, which is a transformed version of that source signal is called deconvolution (for a linear-invariant transform) or deblurring in various fields, such as astronomical image analysis [48]. Whatever the considered application field, the initial learning procedure for estimating the inverse transform may be applied in the nonblind or blind mode, i.e., respectively, with known or unknown input values for the direct system, whereas the output values of that system are known in both modes. This "(unknown) system inversion" task may also be extended to MIMO configurations, in order to restore a set of unknown signals from a set of their transformed versions.

The blind MIMO system inversion problem is also closely related to the field of BSS [30]–[36], whose quantum extension is one of the major topics tackled in this article, together with quantum extensions of system identification. In BSS, the goal is also to restore a set of unknown source signals from a set of available combinations of these signals (called mixtures in BSS), which result from an unknown transform, which combines (i.e., mixes) these source signals. However, in BSS, one most often allows each restored signal to be equal to a source signal only up to an acceptable residual transform (called an indeterminacy) because such transforms cannot be avoided due to the limited constraints that are set on the considered classes of signals and mixing transforms. When applied to the separation of acoustic/audio signals from their mixtures recorded by a set of microphones, BSS is often referred to as the "cocktail party problem" [49].

The first class of BSS methods that was developed and that is still of major importance is ICA [32]–[35]. ICA is a statistical approach, which essentially requires statistically independent random source signals. Thus, ICA is guaranteed to restore the source signals up to limited indeterminacies for the simplest class of mixtures, that is when the available signals are linear instantaneous (i.e., memoryless) combinations of the unknown source signals [32]–[35]. For such mixtures, ICA may be seen as an extension of more conventional PCA [37], [38].

PCA and ICA may both be used to perform mappings from the available P variables to P output variables that are linear instantaneous mixtures of these available variables. In other words, they yield a representation of the same data in a new basis. The selected bases are different in PCA and ICA. PCA uses one of the bases that are such that the output variables are uncorrelated. ICA uses one of the bases that are such that these output variables are statistically independent, which includes uncorrelatedness but is more constraining (for non-Gaussian signals). This is the reason why PCA alone cannot achieve BSS [35] but is often used as a first stage in ICA algorithms. Outside the framework of ICA, PCA is most often used as a mapping that projects the available data onto a lower dimensional space, i.e., with dimension D lower than P, by keeping only the first D coordinates in the output basis, for visualization (with $D = 2$, i.e., projection onto a plane, or $D = 3$, i.e., 3-D visualization) or compression tasks. Such a projection may also be used as a preprocessing stage of ICA, in order to reduce the influence of noise, when the available mixed signals contain noise and their number P is higher than the number M of source signals: One then keeps the first $D = M$ output components of PCA.

ICA also has connections with the above fields of classification and regression in the sense that a significant part of the algorithms developed in all these fields are based on the same class of tools, namely, neural networks. More precisely, when initially developing ICA methods for linear instantaneous mixtures, one of the very first proposed approaches was the well-known Hérault–Jutten neural network (see, e.g., [74]–[78] for its definition and analysis), and extended versions of that network were then introduced and analyzed (see, e.g., [79] and [80]). Neural approaches were then proposed for specific classes of *nonlinear* mixtures or without considering any restrictions on the type of mixture (see, e.g., [81]–[84]). Finally, the interest in neural methods recently raised again also in the field of BSS/ICA. For instance, generative adversarial networks were used to perform linear and nonlinear ICA [85].

The above connected fields of classical system identification, system inversion, BSS, and PCA have been partly extended to the quantum framework as follows. Among these problems, the one that was first studied is the quantum version of nonblind system identification, especially[3] introduced in 1997 in [86] and called QPT by the QIP community (see, e.g., [14]–[23]). The connection between nonblind system identification and regression (and hence, to a lower extent, classification), that we highlighted above, was by the way mentioned for the quantum framework in [65], which states "quantum process tomography (QPT) is able to learn an unknown function within well-defined symmetry and physical constraints—this is useful for regression analysis" and further considers "Regression based on quantum process tomography."

The quantum version of the aforementioned classical source separation, called QSS, and especially its blind version, or BQSS, were then introduced in 2007 in [40]. Two main classes of BQSS methods have been developed since then. The first one may be seen as a quantum extension of the aforementioned classical ICA methods, since it takes advantage of the statistical independence of the parameters

[3]See also [14, p. 398] for the other earliest references.

that define random-coefficient source quantum states (qubit states). It is called quantum independent component analysis (see, e.g., [40] and [41]) or, more precisely, quantum-source independent component analysis (see, e.g., [42]) to insist on the quantum nature of the considered source data, whereas it uses classical processing means (after quantum/classical data conversion). The second main class of BQSS methods was introduced in 2013–2014 in [43] and [44] and then especially detailed in [45]. It is based on the unentanglement of the considered source quantum states, and it typically uses quantum processing means to restore these unknown states from their coupled version. Independently from the above quantum extensions of BSS/ICA, a quantum version of PCA was introduced in 2014 in [39]. Finally, the blind extension of QPT was introduced in 2015 in [24] and then especially extended in [25] for its multiple-preparation version.

Indeed, the above blind or nonblind QPT and (B)QSS methods are restricted to "multiple-preparation" operation, as defined in Section I. Beyond these approaches, this article describes the general SIPQIP framework that may be built to obtain a more efficient operation and its application to various QIP tasks.

APPENDIX B
MULTIPLE-PREPARATION QIP

We here consider an arbitrary number Q of distinguishable [45] qubits, physically implemented as spins 1/2. We investigate the case when the quantum state $|\psi\rangle$ of this set of qubits at a given time is pure and deterministic. As detailed in [26] and [50], the state $|\psi\rangle$ then belongs to a 2^Q-dimensional space $\mathcal{E}$. The vectors that form the standard basis of $\mathcal{E}$ are hereafter denoted as $|k\rangle$, with $k \in \{1, \ldots, 2^Q\}$. The state $|\psi\rangle$ then reads

$$|\psi\rangle = \sum_{k=1}^{2^Q} c_k |k\rangle \tag{27}$$

where c_k are complex-valued coefficients, which are fixed and arbitrary but such that $|\psi\rangle$ is normalized [26], [50].

We here consider the experiment that consists of simultaneously measuring the spin components of all Q qubits along the quantization axis. This experiment yields a random result. Each of its elementary events, [52] A_k may be defined as follows. The result of the experiment is equal to the kth Q-entry vector in the series of 2^Q possible values (in normalized units) equal to $[+\frac{1}{2}, +\frac{1}{2}, \ldots, +\frac{1}{2}, +\frac{1}{2}]$, $[+\frac{1}{2}, +\frac{1}{2}, \ldots, +\frac{1}{2}, -\frac{1}{2}]$ and so on, where these values are, respectively, associated with the above-defined basis vectors $|k\rangle$. Moreover, the probabilities of these events read

$$P(A_k) = |c_k|^2 \qquad \forall\, k \in \{1, \ldots, 2^Q\}. \tag{28}$$

In practice, to estimate the above probabilities for a given Q-qubit state, one most often uses the procedure that will now be defined. That procedure requires one to prepare a large number of copies of that state, which is typically from a few thousand up to a few hundred thousand copies [41], [45]. Therefore we hereafter, call this standard approach "multiple-preparation QIP."[4] These copies may, e.g., be obtained successively from the same system, by using the "repeated write/read" procedure [40]–[42]. The above type of measurement is first performed for each of these copies. One then counts the number of occurrences of each of the possible results $[+\frac{1}{2}, +\frac{1}{2}, \ldots, +\frac{1}{2}, +\frac{1}{2}]$ and so on. The associated sample relative frequencies are then used as estimates of the probabilities $P(A_k)$.

APPENDIX C
BLIND QUANTUM PROCESS TOMOGRAPHY

Various papers from the literature dealing with conventional (i.e., nonblind and multiple-preparation) QPT are focused on specific processes or classes of processes (see, e.g., [19], [21], and [87]). Similarly, the method that we introduced in [26] and that we aim at summarizing in this appendix is dedicated to the class of configurations involving two distinguishable [45] qubits implemented as electron spins 1/2, which are internally coupled according to the cylindrical-symmetry Heisenberg model, with unknown principal values J_{xy} and J_z of the exchange tensor. We stress that this type of coupling is only used as a concrete example,[5] to show how to fully implement the proposed general concepts in a relevant case, but that these concepts and resulting practical algorithms (for performing BQPT and other QIP tasks detailed further in this article) may then be extended to other classes of quantum processes and associated applications.

The above Heisenberg model is detailed in [26]. This shows that the associated quantum process, from its input (i.e., initial) quantum state $|\psi(t_0)\rangle$ to its output (i.e., final) quantum state $|\psi(t)\rangle$, is represented by a matrix M, and that the only quantities that must be estimated in order to obtain an estimate of M are $\exp[i\frac{J_{xy}(t-t_0)}{\hbar}]$ and $\exp[i\frac{J_z(t-t_0)}{2\hbar}]$. The main method proposed in [26] to estimate M uses three values of the time interval $(t - t_0)$, denoted as τ_1, τ_2, and τ_3, with

$$\tau_2 = 2\tau_1 \qquad \tau_3 = 2\tau_2. \tag{29}$$

These values are, respectively, used to first estimate $\exp[i\frac{J_{xy}\tau_1}{\hbar}]$, then estimate $\exp[i\frac{J_z\tau_2}{2\hbar}]$, and finally obtain an estimate of M, which is nonambiguous only from the point of view of the final use of this process with $(t - t_0) = \tau_3$ (Deville and Deville [26] discuss the relevance of finally using a quantum process in conditions, i.e., here with a value of $(t - t_0)$, different from those initially used to identify that process, e.g., when that process corresponds to a gate of a quamputer).

[4]This terminology and the connection between this approach and classical adaptive processing are discussed in [50]. This is also true for the approach of Section II-A.

[5]We do not focus on whether Heisenberg coupling could be used as a desired phenomenon, to build suitable gates for quamputers, as already mentioned in [26].

More precisely, when applying the considered BQPT method in the purely single-preparation mode, the first part of this method uses one instance of each output quantum state $|\psi(t)\rangle$. For each such state, it measures the components of the considered two spins along the Oz-axis. As discussed in [26] and in Appendix B of the present article, the result of each such measurement has four possible values, i.e., $(+\frac{1}{2}, +\frac{1}{2})$, $(+\frac{1}{2}, -\frac{1}{2})$, $(-\frac{1}{2}, +\frac{1}{2})$, or $(-\frac{1}{2}, -\frac{1}{2})$ in normalized units. Their probabilities are, respectively, denoted as p_{1zz} to p_{4zz} below. They are hereafter expressed with respect to the moduli r_j and the phases θ_j and ϕ_j of the polar representation of the qubit parameters α_j and β_j of the single-qubit states

$$|\psi_j(t_0)\rangle = \alpha_j|+\rangle + \beta_j|-\rangle \tag{30}$$

that define the overall input state $|\psi(t_0)\rangle$, which reads

$$|\psi(t_0)\rangle = |\psi_1(t_0)\rangle \otimes |\psi_2(t_0)\rangle \tag{31}$$

$$= \alpha_1\alpha_2|++\rangle + \alpha_1\beta_2|+-\rangle + \beta_1\alpha_2|-+\rangle + \beta_1\beta_2|--\rangle \tag{32}$$

in the 4-D basis $\mathcal{B}_+ = \{|++\rangle, |+-\rangle, |-+\rangle, |--\rangle\}$, and where $\otimes$ stands for the tensor product. The aforementioned polar representation is then defined by

$$\alpha_j = r_j e^{i\theta_j} \qquad \beta_j = q_j e^{i\phi_j}, \qquad j \in \{1, 2\} \tag{33}$$

with

$$q_j = \sqrt{1 - r_j^2}, \qquad j \in \{1, 2\}. \tag{34}$$

The above probabilities then read [41], [42]

$$p_{1zz} = r_1^2 r_2^2 \tag{35}$$

$$p_{2zz} = r_1^2(1 - r_2^2)(1 - v^2) + (1 - r_1^2)r_2^2 v^2 - 2r_1 r_2\sqrt{1 - r_1^2}\sqrt{1 - r_2^2}\sqrt{1 - v^2}\, v \sin\Delta_I \tag{36}$$

$$p_{4zz} = (1 - r_1^2)(1 - r_2^2) \tag{37}$$

with

$$\Delta_I = (\phi_2 - \theta_2) - (\phi_1 - \theta_1) \tag{38}$$

$$\Delta_E = -\frac{J_{xy}(t - t_0)}{\hbar} \tag{39}$$

$$v = \mathrm{sgn}(\cos\Delta_E)\sin\Delta_E. \tag{40}$$

Probability p_{3zz} is not considered hereafter because the sum of p_{1zz} to p_{4zz} is equal to 1.

Using $(t - t_0) = \tau_1$ in the first part of this method, (39) and (40) may then be inverted as

$$\frac{J_{xy}\tau_1}{\hbar} = -\Delta_{Ed} + k_{xy}\pi \tag{41}$$

with

$$\Delta_{Ed} = \arcsin(v) \tag{42}$$

where Δ_{Ed} is a determination associated with the actual value Δ_E, i.e., Δ_{Ed} is equal to Δ_E up to the additive constant $-k_{xy}\pi$, where k_{xy} is an integer.

The SIPQIP framework defined in Section II-A then makes it possible to derive an estimate $\widehat{\Delta}_{Ed}$ of Δ_{Ed} as follows. We consider the case when r_1, r_2, and Δ_I are random valued and when these random variables are statistically independent. Equation (36) then yields

$$E\{p_{2zz}\} = E\{r_1^2\}(1 - E\{r_2^2\})(1 - v^2) + (1 - E\{r_1^2\})E\{r_2^2\}v^2 - 2E\{r_1\sqrt{1 - r_1^2}\}E\{r_2\sqrt{1 - r_2^2}\}\sqrt{1 - v^2}\, v \times E\{\sin\Delta_I\}. \tag{43}$$

In this equation, $E\{p_{2zz}\}$ is known. In practice, it is estimated by using the SIPQIP approach of Section II-A, i.e., by using the sample mean of the estimates of all values of p_{2zz}, themselves typically estimated with sample frequencies (possibly each reduced to one measurement outcome). Similarly, $E\{r_1^2\}$ and $E\{r_2^2\}$ are known: As detailed in [26], they may be derived by solving the two equations obtained by taking the expectation of (35) and (37), which involve $E\{p_{1zz}\}$ and $E\{p_{4zz}\}$, that are also estimated with the SIPQIP approach. Finally, the blind version of QPT concerns the case when the individual values of the input quantum states of the considered process are unknown, but it allows one to request some of the statistical parameters of these inputs to be known. Therefore, we here request the states $|\psi(t_0)\rangle$ to be prepared with a procedure, which is such that the value of $E\{\sin\Delta_I\}$, or at least its sign, is known. Thus, (43) can be exploited so that the only unknown is v. Paper [26] shows how to solve this equation. More precisely, two instances of this equation, with different values of $E\{\sin\Delta_I\}$, are used: The first one yields an estimate of the absolute value of v and the second equation provides an estimate of the sign of v. Combining these two results yields an estimate $\widehat{v}$ of v and, hence, an estimate $\widehat{\Delta}_{Ed}$ of Δ_{Ed} by using $\widehat{v}$ in (42).

Based on (41), once the above estimate $\widehat{\Delta}_{Ed}$ has been obtained, corresponding shifted estimates of $\frac{J_{xy}\tau_1}{\hbar}$ are derived as

$$\frac{\widehat{J_{xy}}\tau_1}{\hbar} = -\widehat{\Delta}_{Ed} + \widehat{k}_{xy}\pi \tag{44}$$

where $\widehat{k}_{xy}$ is an integer, which corresponds to k_{xy} in (41). The value of $\widehat{k}_{xy}$ has to be selected without knowing the actual value of k_{xy} in the fully blind case considered here, i.e., when no prior information is available about the value of J_{xy}. But this is not an issue from the point of view of the considered BQPT method, because that method is designed so that the obtained estimate of the process matrix M, for $(t - t_0) = \tau_3$, does not depend on the integer value of $\widehat{k}_{xy}$ [26]. The simplest approach, therefore, consists of setting $\widehat{k}_{xy} = 0$ in (44).

Similarly, [26] shows that

$$\frac{J_z\tau_2}{\hbar} = \Delta\Phi_{1,0d} + 2k_z\pi + \frac{J_{xy}\tau_2}{\hbar} + \frac{GB\tau_2}{\hbar} \tag{45}$$

where k_z is an integer. This is used to derive the estimate

$$\frac{\widehat{J_z}\tau_2}{\hbar} = \widehat{\Delta\Phi}_{1,0d} + 2\widehat{k_z}\pi + \frac{\widehat{J_{xy}}\tau_2}{\hbar} + \frac{GB\tau_2}{\hbar} \tag{46}$$

where the SIPQIP framework of Section II-A is again used to obtain an estimate $\widehat{\Delta\Phi}_{1,0d}$ of the quantity $\Delta\Phi_{1,0d}$ that may be derived from the same type of probability expectations $E\{p_{kzz}\}$ as above and from the probability expectations $E\{p_{kxx}\}$ of results of additional measurements of spin components along the Ox axis (see details in [26]). Besides, $\widehat{k_z}$ is an integer whose value has no influence on the final estimate of the process matrix M, and that may, therefore, be set to zero. Moreover, $\frac{\widehat{J_{xy}}\tau_2}{\hbar}$ is equal to twice the value previously computed with (44) and the other parameters have known values.

It should be noted that this QPT method only uses the known outputs of the considered process and general known properties of its inputs, not its input *values*, which are unknown. This is, therefore, indeed a BQPT method (moreover operating in the single-preparation mode). In contrast, nonblind methods are supposed to operate with predefined values of their input states and are, in practice, very sensitive to errors in the preparation of these value, as detailed in [26]. Their performance for actual preparations is, therefore, significantly degraded, whereas the above blind operation yields much better accuracy in the tests reported in [26].

APPENDIX D
METHOD FOR ESTIMATING J_{xy} AND J_z

A. ESTIMATING J_{xy}

We here consider the problem of estimating the Hamiltonian parameter J_{xy}, defined in Section III-A. We focus on the practical situation with estimation errors for Δ_{Ed1} and Δ_{Ed2}, and with a known range for J_{xy}. We, hereafter, show how to exploit this range in such a way that the values of $\widehat{J}_{xy1}$ and $\widehat{J}_{xy2}$, which are the closest to one another are also those which are the closest to J_{xy}. To this end, one takes into account that $\widehat{\Delta}_{Ed1}$ and $\widehat{\Delta}_{Ed2}$ are always in the interval $[-\frac{\pi}{2}, \frac{\pi}{2}]$ [because they are values of the arcsin function: see (42)]. This, together with the known range of possible values of $\widehat{J}_{xy1}$, the known value of τ_{11}, and the corresponding version of (44), defines the range $\{\widehat{k}_{xy1}^{\min}, \ldots, \widehat{k}_{xy1}^{\max}\}$ of integers in which it is guaranteed that $\widehat{k}_{xy1}$ should be selected. Similarly, the value of τ_{12} is to be selected as explained hereafter, and for the application of the procedure with any given value τ_{12}, the integer $\widehat{k}_{xy2}$ should be selected in a known interval $\{\widehat{k}_{xy2}^{\min}, \ldots, \widehat{k}_{xy2}^{\max}\}$. When the estimation errors for Δ_{Ed1} and Δ_{Ed2} remain low enough, the values $\widehat{J}_{xy1}$ and $\widehat{J}_{xy2}$ of the grids, respectively, corresponding to $\Delta k_{xy1} = 0$ and $\Delta k_{xy2} = 0$ both remain close to their theoretical value J_{xy}. Around these values, the two grids almost coincide. Then, for larger values of $|\Delta k_{xy1}|$ and $|\Delta k_{xy2}|$ corresponding to the above-defined intervals, we here want the associated parts of the two grids to become more "desynchronized," i.e., we want the gaps between the values of the two grids to become larger. This is obtained by adequately selecting τ_{12} for an arbitrarily chosen value τ_{11}, but this should be performed without knowing where J_{xy} is in the considered interval. We, therefore, use a worst-case approach in terms of desynchronization, for the ideal estimation (10), as follows. The reference point, shared by both grids, is equal to J_{xy} and is obtained when $\Delta k_{xy1} = 0$ and $\Delta k_{xy2} = 0$. We consider the case when this reference point is the lowest value in both bounded grids, i.e., $\widehat{k}_{xy1}^{\min} = k_{xy1}$ and $\widehat{k}_{xy2}^{\min} = k_{xy2}$. For any given τ_{11}, we select a value τ_{12}, which is only somewhat larger than τ_{11}, thus considering that $\widehat{k}_{xy2}^{\min} = \widehat{k}_{xy1}^{\min}$ and $\widehat{k}_{xy2}^{\max} = \widehat{k}_{xy1}^{\max}$. The values are then somewhat closer to one another in the second grid than in the first one. Moreover, we select τ_{12} so that, when moving toward the higher values in both bounded grids, the gaps between the corresponding points of the two grids increase, until they reach the maximum possible gap for the highest values. This means that we set τ_{12} so that the highest value in the first bounded grid [i.e., the value of $\widehat{J}_{xy1}$ in (6) corresponding to $\widehat{k}_{xy1} = \widehat{k}_{xy1}^{\max}$, moreover taking into account (10)] is equal to that of the middle of the interval of the second grid defined as follows: The lower bound of that interval is the highest value in the bounded part of that grid considered here [i.e., the value of $\widehat{J}_{xy2}$ in (8) corresponding to $\widehat{k}_{xy2} = \widehat{k}_{xy2}^{\max} = \widehat{k}_{xy1}^{\max}$, moreover taking into account (10)] and the higher bound of that interval is the next value that would be found in that grid, when moving toward higher values, if that grid were complete, i.e., this upper bound is equal to the lower bound plus $\frac{\hbar\pi}{\tau_{12}}$. Using (6)–(10), it may easily be shown that the above desynchronization condition for the highest values of the two grids yields

$$\frac{\hbar}{\tau_{11}}(\widehat{k}_{xy1}^{\max} - \widehat{k}_{xy1}^{\min})\pi = \frac{\hbar}{\tau_{12}}(\widehat{k}_{xy1}^{\max} - \widehat{k}_{xy1}^{\min} + \frac{1}{2})\pi. \tag{47}$$

Therefore, for a given value τ_{11}, one should set τ_{12} so that

$$\frac{\tau_{12}}{\tau_{11}} = \frac{\widehat{k}_{xy1}^{\max} - \widehat{k}_{xy1}^{\min} + \frac{1}{2}}{\widehat{k}_{xy1}^{\max} - \widehat{k}_{xy1}^{\min}} \tag{48}$$

$$= \frac{2(\widehat{k}_{xy1}^{\max} - \widehat{k}_{xy1}^{\min}) + 1}{2(\widehat{k}_{xy1}^{\max} - \widehat{k}_{xy1}^{\min})}. \tag{49}$$

The latter expression shows that the value thus obtained in this practical procedure for a bounded interval on J_{xy} yields a rational value of $\frac{\tau_{12}}{\tau_{11}}$ (unlike the above preliminary procedure for the ideal case and without restrictions on the domain of J_{xy}).

B. ESTIMATING J_z

The method used for estimating J_z is very similar to the approach described above for J_{xy}. It is, therefore, more briefly outlined hereafter. It uses the procedure of the second part of the BQPT method of Appendix C, based on (45) and

(46). This procedure is here applied twice, i.e., with τ_2 of Appendix C successively replaced by two values denoted as τ_{21} and τ_{22}. For τ_{21}, combining (45) and (46) and using the same type of notations as for J_{xy} yields

$$\widehat{J}_{z1} = J_z + \frac{\hbar}{\tau_{21}}\left[\widehat{\Delta\Phi}_{1,0d1} - \Delta\Phi_{1,0d1} + 2\Delta k_{z1}\pi\right] + (\widehat{J}_{xy} - J_{xy}) \tag{50}$$

with

$$\Delta k_{z1} = \widehat{k}_{z1} - k_{z1} \tag{51}$$

and where $\widehat{J}_{xy}$ is the estimate of J_{xy} without any indeterminacy that was obtained in the first part of this HPE method. This shows that the procedure applied with the time interval τ_{21} yields a regular 1-D grid of possible estimates $\widehat{J}_{z1}$ of J_z, with a step equal to $\frac{2\hbar\pi}{\tau_{21}}$. Its application with the time interval τ_{22} is analyzed in the same way. We here exploit the differences between these two grids, by transposing the approach that we described above for J_{xy}. Thus, first considering the case with no estimation errors and with J_z equal to the lowest value of both bounded grids leads one to select τ_{21} and τ_{22} so that

$$\frac{\hbar}{\tau_{21}}2(\widehat{k}_{z1}^{\max} - \widehat{k}_{z1}^{\min})\pi = \frac{\hbar}{\tau_{22}}2(\widehat{k}_{z1}^{\max} - \widehat{k}_{z1}^{\min} + \frac{1}{2})\pi \tag{52}$$

and hence

$$\frac{\tau_{22}}{\tau_{21}} = \frac{\widehat{k}_{z1}^{\max} - \widehat{k}_{z1}^{\min} + \frac{1}{2}}{\widehat{k}_{z1}^{\max} - \widehat{k}_{z1}^{\min}} \tag{53}$$

where the integers $\widehat{k}_{z1}^{\min}$ and $\widehat{k}_{z1}^{\max}$ are defined by using the same approach as for J_{xy}, here taking into account that $\Delta\Phi_{1,0d1}$, and hence, its relevant estimates are guaranteed to be in the interval $[-\pi, \pi]$ (see the expression of $\Delta\Phi_{1,0d1}$ in [26]) and that $\widehat{J}_{xy}$ and GB are known.

Then, for the practical situation with estimation errors, and still with prior knowledge about an interval that contains the actual value J_z, the method proposed for determining J_z consists of comparing each value $\widehat{J}_{z1}$ of the first bounded grid to each value $\widehat{J}_{z2}$ of the second bounded grid in order to derive the couple of closest values and then the corresponding estimate $\frac{\widehat{J}_{z1} + \widehat{J}_{z2}}{2}$.

APPENDIX E
TEST CONDITIONS

We here define the conditions used for all the tests reported in Section III-A2. The actual values of the parameters of the Heisenberg Hamiltonian defined in [26] were first selected by using the following properties. Conventional electron spin resonance generally operates at X or Q bands (around 10 and 35 GHz, respectively). For electron spins with $g = 2$, at 35 GHz, the resonance field is near 1.25 T. In the simulations, we used the values $g = 2$ and $B = 0.99$ T. Concerning the exchange coupling, we chose $J_z/k_B \simeq 1$ K and $J_{xy}/k_B = 0.3$ K. These values were motivated by [26], [41, Appendix E], and [88]. As in [26], we selected part of the parameters defined above and below so as to avoid specific cases (see [26, footnote 50]), but this here led us to slightly shift some of these values as compared with those of [26], because we here have to take these specific cases into account for *four* time intervals (τ_{11}, τ_{12}, τ_{21}, and τ_{22}) instead of only three (τ_1, τ_2, and τ_3) in [26], so that the BHPE method proposed here is somewhat more constraining that the BQPT method of [26].

The parameters of the BHPE method were then set as follows. The six parameters r_j, θ_j, and ϕ_j, with $j \in \{1, 2\}$, of each initial state $|\psi(t_0)\rangle$ were randomly drawn with a uniform distribution, over an interval that depends on the part of the considered BHPE method, in order to meet the constraints on the statistics of these parameters that are imposed by that BHPE method. The parameters q_1 and q_2 were then derived from (34). More precisely, the parameter J_{xy} was first estimated by applying the procedure of the first part of the BQPT method of Appendix C successively to each of the two values τ_{11} and τ_{12}. For each of these values, as a first step, to estimate the absolute value of v as detailed in [26], the qubit parameter values r_1 and r_2 were selected within the 20–80% subrange of their 0–100% allowed range defined in [26], that is, [0.1, 0.4[for r_1 and [0.6, 0.9[for r_2, as in [42]. Besides, ϕ_1 and ϕ_2 were drawn over $[0, 2\pi[$, whereas θ_1 and θ_2 were fixed to 0 (as stated above, the parameters that have a physical meaning are $\phi_j - \theta_j$). These data are, thus, such that $E\{\sin\Delta_I\} = 0$, as required by this step of the considered BQPT method. Then, as a second step, to estimate the sign of v, as detailed in [26], the same conditions as in the above first step were used for r_j, θ_j, and ϕ_j, with $j \in \{1, 2\}$, except that ϕ_1 was fixed to 0 and ϕ_2 was drawn over $[0, \pi[$. These data are, thus, such that $E\{\sin\Delta_I\}$ is nonzero and has a known sign (here, it is positive), as required by this step of the considered BQPT method. The above two steps were performed with $\tau_{11} = 0.5$ ns and then τ_{12} defined by (48), with $\widehat{k}_{xy1}^{\min} = 0$ and $\widehat{k}_{xy1}^{\max} = 31$ because the only prior knowledge about J_{xy}, which is provided to this BHPE method is that J_{xy}/k_B is in the range [0, 1.5K] (the upper bound 1.5 K was selected as five times the value 0.3 K, which was actually used to create the data processed in these tests as explained above). For τ_{12}, the above interval of values of J_{xy}/k_B results in $\widehat{k}_{xy2}^{\min} = 0$ and $\widehat{k}_{xy2}^{\max} = 32$.

The parameter J_z was then estimated by applying the procedure of the second part of the BQPT method of Appendix C successively to each of the two values τ_{21} and τ_{22}, with $\tau_{21} = 0.53$ ns, and then τ_{22} defined by (53), with $\widehat{k}_{z1}^{\min} = -13$ and $\widehat{k}_{z1}^{\max} = 7$ because the only prior knowledge about J_z which is provided to this BHPE method is that J_z/k_B is in the range [$\simeq$ 0.45K, $\simeq$ 2.24K] (these two bounds were selected as the actual value $\simeq 1$ K, respectively, divided and multiplied by $\sqrt{5}$). For τ_{22}, the above interval of values of J_z/k_B results in $\widehat{k}_{z2}^{\min} = -13$ and $\widehat{k}_{z2}^{\max} = 7$. The proposed method uses measurements along the Oz and Ox axes. For each of the parameters r_j, θ_j, and ϕ_j, with $j \in \{1, 2\}$, we used the same statistics for measurements along the Oz and Ox axes. These

statistics were also the same when using τ_{21} and τ_{22}, and they are defined as follows. The proposed method is based on two instances of (40) of [26]. For the first instance of this equation, r_1 and r_2 were drawn over $[0.1, 0.4[$ and ϕ_1 and ϕ_2 were drawn over $[-\pi/2, \pi/2[$, whereas θ_1 and θ_2 were fixed to 0. For the second instance of the above equation, r_1 and r_2 were drawn over $[0.6, 0.9[$, whereas ϕ_1, ϕ_2, θ_1, and θ_2 were selected in the same way as for the first instance of that equation.

APPENDIX F
COMPUTING THE DOT PRODUCT OF TWO KETS

In Section V, we considered the situation when two known unit-norm classical-form vectors v_1 and v_2 are stored in two unit-norm kets $|\psi_1\rangle$ and $|\psi_2\rangle$, and one then uses quantum circuits from the literature to compute the corresponding overlap $|\langle\psi_1|\psi_2\rangle|^2$. We here propose an extension of this approach, that has not been reported in the literature to our knowledge and that allows one to compute the complex-valued dot product $\langle\psi_1|\psi_2\rangle$ itself, not only its (squared) modulus. To this end, we first consider the classical-form vector

$$v_3 = \mu_3(v_1 + v_2) \tag{54}$$

where μ_3 is real-valued and selected so that v_3 has unit norm. v_3 is stored in the unit-norm ket

$$|\psi_3\rangle = \mu_3(|\psi_1\rangle + |\psi_2\rangle). \tag{55}$$

Simple calculations then yield

$$|\langle\psi_1|\psi_3\rangle|^2 = \mu_3^2(1 + |\langle\psi_1|\psi_2\rangle|^2 + 2\Re(\langle\psi_1|\psi_2\rangle). \tag{56}$$

The aforementioned quantum circuits allow one to compute the overlaps $|\langle\psi_1|\psi_3\rangle|^2$ and $|\langle\psi_1|\psi_2\rangle|^2$, and (56) then yields $\Re(\langle\psi_1|\psi_2\rangle)$. Similarly, $\Im(\langle\psi_1|\psi_2\rangle)$ is obtained by using the ket $|\psi_4\rangle$ corresponding to the vector

$$v_4 = \mu_4(v_1 + iv_2) \tag{57}$$

where μ_4 is real-valued and selected so that v_4 has unit norm. Similar calculations then yield

$$|\langle\psi_1|\psi_4\rangle|^2 = \mu_4^2(1 + |\langle\psi_1|\psi_2\rangle|^2 - 2\Im(\langle\psi_1|\psi_2\rangle). \tag{58}$$

The dot product $>\langle\psi_1|\psi_2\rangle$ is eventually derived from its above real and imaginary parts.

REFERENCES

[1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, 2015, doi: 10.1038/nature14539.
[2] M. Anderson, "Machine learning and quantum computing become BFFs [News]," *IEEE Spectr.*, vol. 54, no. 8, pp. 14–14, Aug. 2017, doi: 10.1109/MSPEC.2017.8000277.
[3] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, "Quantum machine learning," *Nature*, vol. 549, pp. 195–202, Sep. 2017, doi: 10.1038/nature23474.
[4] E. P. DeBenedictis, "A future with quantum machine learning,"*Computer*, vol. 51, no. 2, pp. 68–71, Feb. 2018, doi: 10.1109/MC.2018.1451646.
[5] S. D. Sarma, D.-L. Deng, and L.-M. Duan, "Machine learning meets quantum physics," *Phys. Today*, vol. 72, no. 3, pp. 48–54, Mar. 2019, doi: org/10.1063/PT.3.4164.
[6] C. M. Bishop, *Neural Networks for Pattern Recognition*. Oxford, U.K.: Clarendon Press, 1995.
[7] C. M. Bishop, *Pattern Recognition and Machine Learning*. Singapore: Springer, 2006.
[8] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. New York, NY, USA: Wiley, 2000.
[9] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, 4th ed. San Diego, CA, USA: Academic, 2009.
[10] P. Rebentrost, M. Mohseni, and S. Lloyd, "Quantum support vector machine for big data classification," *Phys. Rev. Lett.*, vol. 113, 2014, Art. no. 130503, doi: 10.1103/PhysRevLett.113.130503.
[11] K. Abed-Meraim, W. Qiu, and Y. Hua, "Blind system identification," *Proc. IEEE*, vol. 85, no. 8, pp. 1310–1322, Aug. 1997, doi: 10.1109/5.622507.
[12] Z. Ding and Y. Li, *Blind Equalization and Identification*. New York, NY, USA: Marcel Dekker, 2001.
[13] L. Ljung, *System Identification: Theory for the User*. Upper Saddle River, NJ, USA: Prentice-Hall, 1999.
[14] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge, U.K.: Cambridge Univ. Press, 2000.
[15] C. H. Baldwin, A. Kalev, and I. Deutsch, "Quantum process tomography of unitary and near-unitary maps," *Phys. Rev. A*, vol. 90, 2014, Art. no. 012110, doi: 10.1103/PhysRevA.90.012110.
[16] R. Blume-Kohout, J. K. Gamble, E. Nielsen, J. Mizrahi, J. D. Sterk, and P. Maunz, "Robust, self-consistent, closed-form tomography of quantum logic gates on a trapped ion qubit," Oct. 2013, *arXiv:1310.4492*.
[17] M. P. A. Branderhorst, J. Nunn, I. A. Walmsley, and R. L. Kosut, "Simplified quantum process tomography," *New J. Phys.*, vol. 11, 2009, Art. no. 115010, doi: 10.1088/1367-2630/11/11/115010.
[18] S. T. Merkel *et al.*, "Self-consistent quantum process tomography," *Phys. Rev. A*, vol. 87, 2013, Art. no. 062119, doi: 10.1103/PhysRevA.87.062119.
[19] N. Navon, N. Akerman, S. Kotler, Y. Glickman, and R. Ozeri, "Quantum process tomography of a Mølmer–Sørensen interaction," *Phys. Rev. A*, vol. 90, 2014, Art. no. 010103, doi: 10.1103/PhysRevA.90.010103.
[20] A. Shukla and T. S. Mahesh, "Single-scan quantum process tomography," *Phys. Rev. A*, vol. 90, 2014, Art. no. 052301, doi: 10.1103/PhysRevA.90.052301.
[21] M. Takahashi, S. D. Bartlett, and A. C. Doherty, "Tomography of a spin qubit in a double quantum dot," *Phys. Rev. A*, vol. 88, 2013, Art. no. 022120, doi: 10.1103/PhysRevA.88.022120.
[22] Y. Wang, D. Dong, I. R. Petersen, and J. Zhang, "An approximate algorithm for quantum Hamiltonian identification with complexity analysis," in *Proc. 20th World Congr. Int. Federation Autom. Control*, Toulouse, France, Jul. 9–14, 2017, pp. 12241–12245., doi: 10.1016/j.ifacol.2017.08.1949
[23] A. G. White and A. Gilchrist, G. J. Pryde, J. L. O'Brien, M. J. Bremner, and N. K. Langford, "Measuring two-qubit gates," *J. Opt. Soc. Amer. B*, vol. 24, no. 2, pp. 172–183, Feb. 2007, doi: 10.1364/JOSAB.24.000172.
[24] Y. Deville and A. Deville, "From blind quantum source separation to blind quantum process tomography," in *Proc. 12th Int. Conf. Latent Variable Anal. Signal Separation (LVA/ICA 2015)*, Aug. 25–28, 2015, pp. 184–192, doi: 10.1007/978-3-319-22482-4_21.
[25] Y. Deville and A. Deville, "The blind version of quantum process tomography: Operating with unknown input values," in *Proc. 20th World Congr. Int. Federation Autom. Control*, Toulouse, France, Jul. 9–14, 2017, pp. 12228–12234, doi: 10.1016/j.ifacol.2017.08.1947.
[26] Y. Deville and A. Deville, "Quantum process tomography with unknown single-preparation input states: Concepts and application to the qubit pair with internal exchange coupling," *Phys. Rev. A*, vol. 101, no. 4, Apr. 2020, Art. no. 042332, doi: 10.1103/PhysRevA.101.042332.
[27] A. Melkani, C. Gneiting, and F. Nori, "Eigenstate extraction with neural-network tomography," *Phys. Rev. A*, vol. 102, no. 2, 2020, Art. no. 022412, doi: 10.1103/PhysRevA.102.022412.
[28] S. Ahmed, C. S. Munoz, F. Nori, and A. F. Kockum, "Quantum state tomography with conditional generative adversarial networks," *Phys. Rev. Lett.*, vol. 127, 2021, Art. no. 140502, doi: 10.1103/PhysRevLett.127.140502.
[29] S. Ahmed, C. S. Munoz, F. Nori, and A. F. Kockum, "Classification and reconstruction of optical quantum states with deep neural networks," Dec. 2020, *arXiv:2012.02185*.

[30] A. Cichocki and S.-I. Amari, *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*. Chichester, U.K.: Wiley, 2002.

[31] A. Cichocki, R. Zdunek, A. H. Phan, and S.-I. Amari, *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*. Chichester, U.K.: Wiley, 2009.

[32] P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent Component Analysis and Applications*. Oxford, U.K.: Academic, 2010.

[33] Y. Deville, *Traitement Du Signal : Signaux Temporels Et Spatiotemporels—Analyse Des Signaux, Théorie De L'information, Traitement D'antenne, Séparation Aveugle De Sources*. Paris, France: Ellipses Editions Marketing, 2011.

[34] Y. Deville, "Blind source separation and blind mixture identification methods," in *Wiley Encyclopedia of Electrical and Electronics Engineering*, J. Webster Ed. Hoboken, NJ, USA: Wiley, pp. 1–33.

[35] A. Hyvarinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. New York, NY, USA: Wiley, 2001.

[36] S. Makino, H. Sawada, and T.-W. Lee, Eds., *Blind Speech Separation*. Dordrecht, The Netherlands: Springer, 2007.

[37] H. Abdi and L. J. Williams, "Principal component analysis," *Wires Comput. Statist.*, vol. 2, pp. 433–459, Jul./Aug. 2010, doi: 10.1002/wics.101.

[38] I. T. Jolliffe, *Principal Component Analysis*. New York, NY, USA: Springer-Verlag, 2002.

[39] S. Lloyd, M. Mohseni, and P. Rebentrost, "Quantum principal component analysis," *Nature Phys.*, vol. 10, pp. 631–633, Sep. 2014, doi: 10.1038/nphys3029.

[40] Y. Deville and A. Deville, "Blind Separation of quantum states: Estimating two qubits from an isotropic spin coupling model," in *Proc. 7th Int. Conf. Independent Compon. Anal. Signal Separation (ICA 2007)*, Sep. 2007, pp. 706–713, doi: 10.1007/978-3-540-74494-8_88.

[41] Y. Deville and A. Deville, "Classical-processing and quantum-processing signal separation methods for qubit uncoupling," *Quantum Inf. Process.*, vol. 11, no. 6, pp. 1311–1347, 2012, doi: 10.1007/s11128-011-0273-7.

[42] Y. Deville and A. Deville, "Quantum-source independent component analysis and related statistical blind qubit uncoupling methods," in *Blind Source Separation: Advances in Theory, Algorithms and Applications*, G. R. Naik and W. Wang, Eds. Berlin, Germany:Springer, 2014, pp. 3–37.

[43] Y. Deville and A. Deville, "A quantum-feedforward and classical-feedback separating structure adapted with monodirectional measurements; blind qubit uncoupling capability and links with ICA," in *Proc. 23rd IEEE Int. Workshop Mach. Learn. Signal Process.*, 2013, pp. 1–6, doi: 10.1109/MLSP.2013.6661954.

[44] Y. Deville and A. Deville, "Blind qubit state disentanglement with quantum processing: Principle, criterion and algorithm using measurements along two directions," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP 2014)*, Florence, Italy, May 4–9, 2014, pp. 6221–6225, doi: 10.1109/ICASSP.2014.6854800.

[45] Y. Deville and A. Deville, "Blind quantum source separation: Quantum-processing qubit uncoupling systems based on disentanglement," *Digit. Signal Process.*, vol. 67, pp. 30–51, Aug. 2017, doi: 10.1016/j.dsp.2017.04.013.

[46] J. G. Proakis, *Digital Communications*. Boston, MA, USA: McGraw-Hill, 2001.

[47] R. D. Mori, *Spoken Dialogues With Computers*. London, U.K.: Academic, 1998.

[48] J.-L. Starck and F. Murtagh, *Astronomical Image and Data Analysis*. Berlin, Germany: Springer, 2006.

[49] E. C. Cherry, "Some experiments on the recognition of speech, with one and with two ears," *J. Acoust. Soc. Amer.*, vol. 25, no. 5, pp. 975–979, Sep. 1953, doi: 10.1121/1.1907229.

[50] Y. Deville and A. Deville, "Stochastic quantum information processing, with applications to blind quantum system identification and source separation," in *Proc. IEEE 28th Int. Workshop Mach. Learn. Signal Process.*, Aalborg, Denmark, Sep. 17–20, 2018, pp. 1–6, doi: 10.1109/MLSP.2018.8517050.

[51] A. Deville and Y. Deville, "Concepts and criteria for blind quantum source separation and blind quantum process tomography," *Entropy*, vol. 19, no. 7, 2017, Art. no. 311, doi: 10.3390/e19070311.

[52] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. Singapore: McGraw-Hill, 1984.

[53] J. Zhang and M. Sarovar, "Quantum Hamiltonian identification from measurement time traces," *Phys. Rev. Lett.*, vol. 113, 2014, Art. no. 080401, doi: 10.1103/PhysRevLett.113.080401.

[54] A. Cooper, W. K. C. Sun, J.-C. Jaskula, and P. Cappellaro, "Identification and control of electron-nuclear spin defects in diamond," *Phys. Rev. Lett.*, vol. 124, no. 8, Feb. 2020, Art. no. 083602, doi: 10.1103/PhysRevLett.124.083602.

[55] H. Yuan and C.-H. F. Fung, "Optimal feedback scheme and universal time scaling for Hamiltonian parameter estimation," *Phys. Rev. Lett.*, vol. 115, 2015, Art. no. 110401, doi: 10.1103/PhysRevLett.115.110401.

[56] J. Geremia and H. Rabitz, "Optimal identification of Hamiltonian information by closed-loop laser control of quantum systems," *Phys. Rev. Lett.*, vol. 89, no. 26, Dec. 2002, Art. no. 263902, doi: 10.1103/PhysRevLett.89.263902.

[57] L. Tan, D. Dong, D. Li, and S. Xue, "Quantum Hamiltonian identification with classical colored measurement noise," *IEEE Trans. Control Syst. Technol.*, vol. 29, no. 3, pp. 1356–1363, May 2021, doi: 10.1109/TCST.2020.2991611.

[58] Y. Wang, D. Dong, B. Qi, J. Zhang, I. R. Petersen, and H. Yonezawa, "A quantum Hamiltonian identification algorithm: Computational complexity and error analysis," *IEEE Trans. Autom. Control*, vol. 63, no. 5, pp. 1388–1403, May 2018, doi: 10.1109/TAC.2017.2747507.

[59] D. Burgarth, K. Maruyama, and F. Nori, "Coupling strength estimation for spin chains despite restricted access," *Phys. Rev. A*, vol. 79, no. 2, 2009, Art. no. 020305(R), doi: 10.1103/PhysRevA.79.020305.

[60] D. Burgarth *et al.*, "Scalable quantum computation via local control of only two qubits," *Phys. Rev. A*, vol. 81, no. 4, 2010, Art. no. 0 40303(R), doi: 10.1103/PhysRevA.81.040303.

[61] D. Burgarth, K. Maruyama, and F. Nori, "Indirect quantum tomography of quadratic Hamiltonians," *New J. Phys.*, vol. 13, 2011, Art. no. 013019, doi: 10.1088/1367-2630/13/1/013019.

[62] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1975, Art. no. 07632.

[63] J. Preskill, "Lecture notes for PH219/CS219: Quantum information and computation," 2018. [Online]. Available: http://www.theory.caltech.edu/preskill/ph219/chap3_15.pdf

[64] H. Qiao *et al.*, "Coherent multispin exchange coupling in a quantum-dot spin chain," *Phys. Rev. X*, vol. 10, no. 3, 2020, Art. no. 031006, doi: 10.1103/PhysRevX.10.031006.

[65] P. Wittek, Quantum Machine Learning: What Quantum Computing Means to Data Mining. Amsterdam, The Netherlands: Academic, 2014.

[66] F. A. Kruse *et al.*, "The spectral image processing system (SIPS)—Interactive visualization and analysis of imaging spectrometer data," *Remote Sens. Environ.*, vol. 44, pp. 145–163, 1993, doi: 10.1016/0034-4257(93)90013-N.

[67] S. Lloyd, M. Mohseni, and P. Rebentrost, "Quantum algorithms for supervised and unsupervised machine learning," Nov. 2013, *arXiv:1307.0411v2*.

[68] L. Cincio, Y. Subasi, A. T. Sornborger, and P. J. Coles, "Learning the quantum algorithm for state overlap," Nov. 2018, *arXiv:1307.0411*.

[69] A. Peres, *Quantum Theory: Concepts and Methods*. Dordrecht, The Netherlands: Kluwer, 1995.

[70] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf, "Quantum fingerprinting," *Phys. Rev. Lett.*, vol. 87, no. 16, 2001, Art. no. 167902, doi: 10.1103/PhysRevLett.87.167902.

[71] D. H. Johnson and D. Dudgeon, *Array Signal Processing: Concepts and Techniques*. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.

[72] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 1996.

[73] G. R. Steinbrecher, J. P. Olson, D. Englund, and J. Carolan, "Quantum optical neural networks," *Nature Partner J. Quantum Inf.*, vol. 5, no. 60, pp. 1–9, 2019, doi: 10.1038/s41534-019-0174-7.

[74] J. Hérault and B. Ans, "Circuits neuronaux synapses modifiables: Décodage de messages composites par apprentissage non-supervisé," *C. R. de l'Académie des Sciences de Paris*, vol. 299, no. 13, pp. 525–528, 1984.

[75] C. Jutten and J. Hérault, "Blind separation of sources, Part I: An adaptive algorithm based on neuromimetic architecture," *Signal Process.*, vol. 24, no. 1, pp. 1–10, Jul. 1991, doi: 10.1016/0165-1684(91)90079-X.

[76] P. Comon, C. Jutten, and J. Hérault, "Blind separation of sources, Part: II: Problems statement," *Signal Process.*, vol. 24, no. 1, pp. 11–20, 1991, doi: 10.1016/0165-1684(91)90080-3.

[77] E. Sorouchyari, "Blind separation of sources, Part III: Stability analysis," *Signal Process.*, vol. 24, no. 1, pp. 21–29, Jul. 1991, doi: 10.1016/0165-1684(91)90081-S.

[78] Y. Deville, "A unified stability analysis of the Hérault-Jutten source separation neural network," *Signal Process.*, vol. 51, no. 3, pp. 229–233, Jun. 1996, doi: 10.1016/0165-1684(96)00046-1.

[79] A. Cichocki, W. Kasprzak, and S. Amari, "Multi-layer neural networks with a local adaptive learning rule for blind separation of source signals," in *Proc. Int. Symp. Nonlinear Theory Appl.*, Dec. 1995, pp. 61–65.

[80] Y. Deville, "Analysis of the convergence properties of self-normalized source separation neural networks," *IEEE Trans. Signal Process.*, vol. 47, no. 5, pp. 1272–1287, May 1999, doi: 10.1109/78.757215.

[81] L. B. Almeida, "MISEP–Linear and nonlinear ICA based on mutual information," *J. Mach. Learn. Res.*, vol. 4, pp. 1297–1318, 2003, doi: 10.1162/jmlr.2003.4.7-8.1297.

[82] Y. Deville and S. Hosseini, "Recurrent networks for separating extractable-target nonlinear mixtures. Part I: Non-blind configurations," *Signal Process.*, vol. 89, no. 4, pp. 378–393, Apr. 2009, doi: 10.1016/j.sigpro.2008.09.016.

[83] L. Duarte, F. O. Pereira, R. Attux, R. Suyama, and J. M. T. Romano, "Source Separation in post-nonlinear mixtures by means of monotonic networks," in *Proc. 12th Int. Conf. Latent Variable Anal. Signal Separation*, Aug. 2015, pp. 176–183, doi: 10.1007/978-3-319-22482-4_20.

[84] S. Hosseini and Y. Deville, "Recurrent networks for separating extractable-target nonlinear mixtures. Part II: Blind configurations," *Signal Process.*, vol. 93, no. 4, pp. 671–683, Apr. 2013, doi: 10.1016/j.sigpro.2012.08.027.

[85] P. Brakel and Y. Bengio, "Learning Independent features with adversarial nets for non-linear ICA," Oct. 2017, *arXiv:1710.05050*.

[86] I. L. Chuang and M. A. Nielsen, "Prescription for experimental determination of the dynamics of a quantum black box," *J. Modern Opt.*, vol. 44, nos. 11/12, pp. 2455–2467, 1997, doi: 10.1080/09500349708231894.

[87] J. L. O'Brien *et al.*, "Quantum process tomography of a controlled-not gate," *Phys. Rev. Lett.*, vol. 93, no. 8, 2004, Art. no. 080502, doi: 10.1103/PhysRevLett.93.080502.

[88] A. Ferretti, M. Fanciulli, A. Ponti, and A. Schweiger, "Electron spin-echo relaxation and envelope modulation of shallow phosphorus donors in silicon," *Phys. Rev. B*, vol. 72, 2005, Art. no. 235201, doi: 10.1103/PhysRevB.72.235201.

Yannick Deville (Member, IEEE) received an Engineering degree in telecommunications from the Ecole Nationale Supérieure des Télécommunications de Bretagne, Brest, France, in 1986, and the D.E.A. and Ph.D. degrees in microelectronics from the University of Grenoble, Grenoble, France, in 1986 and 1989, respectively.

From 1986 to 1997, he was with Philips Research Labs, Limeil, France. Since 1997, he has been a Full Professor with the University of Toulouse, Toulouse, France. His research interests include signal/image processing, higher order statistics, time–frequency analysis, neural networks, quantum entanglement, and especially blind source separation and blind identification methods and their applications to remote sensing, astrophysics, and quantum information processing.

Alain Deville is presently a Professor Emeritus at Aix-Marseille Université, Marseille, France. After his Agrégation de Physique (1967), and two years spent at the Université de Poitiers, in 1969 he joined the Physics Department at the Université de Provence in Marseille, where he contributed to the development of the teaching of electronics. He received his Thèse d'Etat in 1976.

He was a full Professor in 1984, and has been a Professor Emeritus since 2002. He has worked both experimentally (development and use of microwave equipments for continuous wave and transient ESR) and theoretically on the microscopic equilibrium and relaxation magnetic properties of electron spins in solids (ionic insulators, glasses and spin glasses, high Tc superconductors). During the 2000–2005 years, with his wife Danielle Deville, he wrote a three-volume treatise devoted to the physics of electronic components and circuits, for Students of Electronics. Since 2005, he has been involved in problems related to quantum information processing.

Prof. Deville is a member of the American Physical Society.