Universality and optimality of programmable quantum processors

Mário Ziman[1,2,3] and Vladimír Bužek[1,2,3]
[1] *Research Center for Quantum Information,*
Slovak Academy of Sciences,
Dúbravská cesta 9, 845 11 Bratislava, Slovakia
[2] Quniverse, *Líščie údolie 116,*
841 04 Bratislava, Slovakia
[3] *Faculty of Informatics, Masaryk University,*
Botanická 68a, 602 00 Brno, Czech Republic

We analyze and compare the optimality of approximate and probabilistic universal programmable quantum processors. We define several characteristics how to quantify the optimality and we study in detail performance of three types of programmable quantum processors based on (1) the C-NOT gate, (2) the SWAP operation, and (3) the model of the quantum information distributor - the QID processor. We show under which conditions the measurement assisted QID processor is optimal. We also investigate optimality of the so-called U-processors and we also compare the optimal approximative implementation of U(1) qubit rotations with the known probabilistic implementation as introduced by Vidal, Masanes and Cirac [*Phys. Rev. Lett.* **88**, 047905 (2002)].

I. PROGRAMMABLE PROCESSORS

Classical programmable processors are realized as a hardware that perform an operation (called computation) on a data register according to instructions (program) encoded in a program register (software). It is one of the central issues of computer science whether there exist a universal (classical) processor that performs all possible classical transformations of the data register of the size of d bits [1]. Let us, for simplicity, consider only reversible classical computation (though the conclusions are valid also for irreversible classical programs). The register composed of d classical bits can be reversibly transformed in $2^d!$ different ways (permutations), e.g. we consider that on a single bit we apply only two programs: the identity and the NOT operation. For a single-bit data register the controlled-NOT (CNOT) gate serves as a universal classical processor [the CNOT gate is defined as follows CNOT : $j,k \to j, j\oplus k$ $(j,k=0,1)$]. That is, if the program register (consisting of a single bit) has the value (is in the state) $k=0$ then the identity operation is realized on the data bit. Similarly, if $k=1$ then the NOT operation is performed on the data bit. It is clear that such "control" devices realizing different programs for different bit values of the program register, are universal programmable processors also for larger data registers. The size of the program register (in terms of the number N of bits) is given by the relation $2^d! = 2^N$, because 2^N represents the number of different states of the register of the size N. As a result we obtain that universal processor for d-bit data register consists of approximately $N \approx 2^d(d-1)$ bits, i.e. it is exponentially large.

In their seminal paper [2] Nielsen and Chuang showed that the *quantum* analogue of universal programmable processor does not exist, i.e. it is not possible even in principle to design a universal deterministic programmable quantum processor. Information about the quantum operation cannot be encoded into the state of arbitrarily large program register. In order to realize n unitary transformation of the data register one must use n dimensional program register. The same holds also in the classical case, but unlike there, in the quantum case even for a single qubit the number of possible programs (unitary transformations) is uncountably infinite. This requires inseparable quantum systems as program registers, but such systems are usually excluded from the standard quantum theory. The nonexistence of universal programmable quantum device is another example of no-go theorems in quantum information processing.

Let us consider two completely positive maps $\mathcal{E},\mathcal{F}$ given by Kraus operators $\{E_j\},\{F_k\}$, respectively. In addition, let us assume that these two operations can be realized with a fixed processor G, i.e. a unitary operation acting on the joint data plus program Hilbert space $\mathcal{H}_d\otimes\mathcal{H}_p$. Denote $|\Xi_E\rangle, |\Xi_F\rangle$ the corresponding pure states of the program register, i.e. $\mathcal{E}[\varrho]=\sum_j E_j\varrho E_j^\dagger$ with $E_j=\langle j|G|\Xi_E\rangle$, where $\mathcal{F}[\varrho]=\sum_j F_j\varrho F_j^\dagger$ with $F_j=\langle j|G|\Xi_F\rangle$, and $\{|j\rangle\}$ is some fixed basis of $\mathcal{H}_p$. Calculating $\sum_j E_j^\dagger F_j$ one derives the following identity

$$\sum_j E_j^\dagger F_j = \langle\Xi_E|\Xi_F\rangle I = cI\,. \tag{1.1}$$

This equation is necessary for simultaneous realization of both operations $\mathcal{E},\mathcal{F}$ on the processor G. We remind us the ambiguity of Kraus decomposition, i.e. when applying this criterion one must take into account all possible decompositions. A special case is achieved when $c = 0$, i.e. the encoding program states are orthogonal. In such a way any finite number of quantum operations can be realized by encoding them into mutually orthogonal states. Let us apply this condition to the case of unitary operations $U_1, \ldots, U_n$. For them the Kraus decomposition is unique and we obtain the set of conditions $U_j^\dagger U_k = c_{jk} I$. Obviously $c_{jj} = 1$, but in all other cases ($j \neq k$) it is necessary to set $c_{jk} = 0$ in order to satisfy Eq. (1.1). It means that encoding n unitaries requires n-dimensional program space. By construction it can be seen that such dimension of the program Hilbert space is also sufficient. One can simply define the processor as $G = \sum_j U_j \otimes |j\rangle\langle j|$. Using this criterion of compatibility (1.1) one can investigate the programmability of different families of processes. For instance, in Ref. [3] we have shown that the family of phase damping channels can be implemented on a quantum processor, but the family of amplitude damping channels cannot.

Properties of quantum programmable processors have been studied already by many authors and from different perspectives [4–10]. In the present paper we focus our attention on optimality and universality of approximate and probabilistic programmable quantum processors. The paper is organized as follows: In the Section II we study optimality of programmable quantum processor, in the Section III we analyze in detail optimality of three models of programmable quantum processors: the CNOT, the QID and the SWAP processors. The Section IV is devoted to approximative programming. We will show in which sense the QID processors are optimal. In the Section V we relax the universality condition and we discuss programmable processor that allows us to implement one-parametric group of single-qubit rotations. We conclude our paper in Section VI with a brief summary of our results.

II. UNIVERSALITY AND OPTIMALITY

Even though universal deterministic programmable quantum processors do not exist [2] one can investigate various approximations of these processors (this is a general approach when one deals with quantum-mechanical no-go theorems). One can study scenarios how to achieve the universality by relaxing few of the ideal conditions. In principle, there are two options: i) an approximative implementation of quantum programs, or ii) a probabilistic implementation of quantum programs. In the first case we allow some imprecision ϵ in the implementation of the desired quantum programs, whereas in the second case we relax the condition that the programmability is deterministic by introducing a concept of the success probability P_{success}. Important point is that even though the implementation of the program is only probabilistic, the measurement outcomes tell us exactly when the desired operation is performed. It is not difficult to see that in both of these cases the *universal* programmable processors (either approximate or probabilistic) do exist [4, 6, 11–13]. However, since the universality is conditioned by some imperfections the question of optimality of encoding of quantum operations into states of quantum program registers is of importance [6, 13, 14].

For a given processor G one can should study whether it is universal in an approximate, or in a probabilistic sense (or both). Different (approximate or probabilistic) universal programmable processors can be compared with the help of the approximation parameter ϵ, or the probability success P_{success}, respectively. There are several ways how to characterize the optimality via these parameters. Let remind us the exact definition of these parameters:

The approximate programmable processors perform a quantum operation $\mathcal{T}$ with the precision $\epsilon(\mathcal{T}) = \min_\xi D(\mathcal{T}, \mathcal{E}_\xi)$, where $\mathcal{E}_\xi[\varrho] = \mathrm{Tr}_p G(\varrho \otimes \xi) G^\dagger$ and $D(.,.)$ quantifies the distance between two quantum operations. In the probabilistic case we perform a measurement on the program register at the output of the processor. In principle, we can distinguish two cases: either the measurement is fixed (measurement-assisted processor), or the choice of an arbitrary von Neumann measurement is a part of the quantum programming. After recording an outcome m of the measurement of the program register the data register is transformed into the state $\varrho_m = \mathcal{T}_m[\varrho] = \frac{1}{p_m}\mathcal{I}_m[\varrho]$, where $p_m = \mathrm{Tr}\mathcal{I}_m[\varrho]$ and $\mathcal{I}_m$ is a linear completely positive, but not necessarily trace-preserving, map. Only in cases when $p_m \neq p_m(\varrho)$ the transformation $\mathcal{T}_m$ corresponds to some quantum operation, i.e. it is a completely positive trace-preserving linear map. Without the measurement, or better to say without the post-selection, the data register is transformed by some quantum operation $\mathcal{E}_\xi$. These maps are always performed with a probability equal to unity. For each program state one can express the realized operation as a convex combination $\mathcal{E}_\xi = \sum_j q_j \mathcal{T}_j$ and the problem is whether it is possible to find a measurement M such that the operation $\mathcal{T}_j$ was realized with the probability q_j. The decomposition of $\mathcal{E}_\xi = \sum_m q_m \mathcal{T}_m$ is called realizable if there exists a measurement M of the program register with outcomes m such that $\mathcal{T}_m = \frac{1}{p_m}\mathcal{I}_m$. The success probability $P_{\text{success}}(\mathcal{T})$ of the operation $\mathcal{T}$ is defined as the maximum of probabilities p over all program states ξ with the realizable decomposition $\mathcal{E}_\xi = p\mathcal{T} + (1-p)\mathcal{N}$. The situation is simple if one uses the measurement-assisted quantum processor, i.e. the measurement is fixed for all inputs. The universality of such device was demonstrated explicitly in Ref. [11], but the optimality is still an open question. It is clear that limits of measurement-assisted processors are stronger than limits for probabilistic processors where one can vary measurements in order to increase success probabilities.

There are several approaches how to compare performance of quantum programmable processors. One can use either extremal (worst) cases, or average values to evaluate the accuracy/success of approximate/probabilistic processors, i.e.

$$\overline{P}^{G}_{\rm success}=\int_{\mathcal{T}}{\rm d}\mathcal{T}P_{\rm success}(\mathcal{T}),\quad \overline{\epsilon}_G=\int_{\mathcal{T}}{\rm d}\mathcal{T}\epsilon(\mathcal{T})$$
$$P^{G}_{\rm success}=\min_{\mathcal{T}}P_{\rm success}(\mathcal{T}),\quad \epsilon_G=\max_{\mathcal{T}}\epsilon(\mathcal{T})$$

There is also a freedom in the choice of the function $D(.,.)$ in the definition of the accuracy $\epsilon(\mathcal{T})$ of the approximation. The larger the success probability the better the probabilistic processor and similarly, the smaller the error the better the approximate processor. We will use the notation $P^{G,M}_{\rm success}$ and $\overline{P}^{G,M}_{\rm success}$ for the parameters of the measurement-assisted quantum processor specified by the unitary transformation G and the program-register measurement M. The following relations hold: $\max_M P^{G,M}_{\rm success}\leq P^{G}_{\rm success}$ and $\max_M \overline{P}^{G,M}_{\rm success}\leq \overline{P}^{G}_{\rm success}$.

Also the universality of quantum processors is usually understood in two different ways: i) either with respect to an implementation of all quantum operations, or ii) with respect to a realization of all unitary transformations of the program register. In some cases the set of implemented operations can be reduced (restricted) to smaller families of quantum operations. It is clear that due to a unitary representation of any completely positive tracepreserving linear map both of the meaning of universality are closely related. In order to define optimal quantum processor we usually fix the size of the data and program register, $d={\rm dim}\mathcal{H}_d$ and $N={\rm dim}\mathcal{H}_p$. The main open problem of quantum processor's optimality is to find this functional dependence for success probabilities and approximation accuracy. In other words, for a fixed length of data and program register the task is to find the class of processors maximizing the success probability and minimizing the approximation parameter.

Let us take a process fidelity to quantify the distance between two maps, i.e. $D(\mathcal{E},\mathcal{T})=1-F(\mathcal{E},\mathcal{T})=1-f(\Phi_{\mathcal{E}},\Phi_{\mathcal{T}})$, where $\Phi_{\mathcal{E}}=\mathcal{E}\otimes\mathcal{I}[\Psi_+]$ (Ψ_+ is the projector onto maximally entangled state $|\psi_+\rangle=\frac{1}{\sqrt{d}}\sum_j|j\rangle\otimes|j\rangle$) and $f(\varrho,\sigma)=({\rm Tr}\sqrt{\sqrt{\varrho}\sigma\sqrt{\varrho}})^2$ is the state fidelity function. From the definition of the quantities $P_{\rm success}(\mathcal{T})$, $\epsilon(\mathcal{T})$ and properties of process fidelity it follows that [13]

$$P_{\rm error}(\mathcal{T})\geq\epsilon(\mathcal{T})\,, \tag{2.1}$$

where $P_{\rm error}(\mathcal{T})=1-P_{\rm success}(\mathcal{T})$. The probabilistic realization of $\mathcal{T}$ means that a given program state ξ induces the map $\mathcal{E}_\xi=P_{\rm success}\mathcal{T}+P_{\rm error}\mathcal{N}$. One can say that program encoded in state ξ approximates the transformation $\mathcal{T}$ with the precision quantified by $\epsilon=1-F(\mathcal{E}_\xi,\mathcal{T})$. Using the concavity of process fidelity one can directly show that $\epsilon\leq 1-P_{\rm success}+P_{\rm error}F(\mathcal{N},\mathcal{T})\leq 1-P_{\rm success}$ which gives the above inequality.

Similar relations hold also for the derived average, and worst case quantities. This inequality is not saturated in general (for fixed quantum processor), however it might be the case that for the optimal values (optimized over all processors) these two numbers coincide. It could be an interesting result if, moreover, the same optimal processors are optimal for approximate as well as probabilistic scenario.

For our purposes it will be useful to have an expression for the process fidelity between an arbitrary channel $\mathcal{E}$ (defined as $\mathcal{E}[\varrho]=\sum_r A_r\varrho A_r^\dagger$) and a unitary transformation U. The state $\Phi_U=U\otimes I[\Psi_+]$ is pure, i.e. $\sqrt{\Phi_U}=\Phi_U$. It follows that

$$F(U,\mathcal{E})=\langle\Phi_U|\mathcal{E}\otimes\mathcal{I}[\Psi_+]|\Phi_E\rangle=\frac{1}{d^2}\sum_{j,k}\langle j|U^\dagger\mathcal{E}[|j\rangle\langle k|]U|k\rangle=\frac{1}{d^2}\sum_r|{\rm Tr}U^\dagger A_r|^2\,.$$

III. PROGRAMMABILITY OF UNITARY TRANSFORMATIONS

In this section we will pay attention to a simpler problem of implementation of all unitary maps. In such case there exists a unique Haar measure dU that enables us, in principle, to calculate also the average error and average success probability. In what follows we shall analyze three examples of quantum processors.

A. Controlled NOT

At the beginning of this paper we have seen that the controlled NOT (CNOT) gate serves as universal classical programmable processor implementing the programs on a single classical bit. Its quantum version, $G_{\rm CNOT}=I\otimes|+\rangle\langle+|+\sigma_z\otimes|-\rangle\langle-|$ (with $|\pm\rangle=\frac{1}{\sqrt{2}}(|0\rangle\pm|1\rangle)$), is not universal, but still it can serve as a good and simple example of a quantum processor realizing approximatively and probabilistically a specific subclass of unitary operations. Using this device we are able to realize deterministically arbitrary channel of the form $\mathcal{E}[\varrho]=p\varrho+(1-p)\sigma_z\varrho\sigma_z$. Vidal and Cirac [15, 16] showed that using the measurement-assisted CNOT processor one can probabilistically implement arbitrary $U(1)$ rotation $U_\varphi = \exp(-i\varphi\sigma_z)$ with the probability $P_{\rm success}(U_\varphi) = 1/2$. The fixed measurement is specified by the basis $|0\rangle, |1\rangle$.

In Ref. [17] we studied different choices of measurement for the CNOT processor. We found that arbitrary measurement enables us to realize probabilistically a whole set of unitary transformations U_φ, though in this case probability distributions $P^M_{\rm success}(U_\varphi)$ are different for different measurements M. Of course, the transformations I, σ_z are always implemented with the probability equal to. The success probability for a fixed measurement M is given by the relation $P_{success} = \cos^2\xi\cos^2\eta + \sin^2\xi\sin^2\eta$, where the angle η specifies the choice of the measurement and ξ represents the states encoding the unitary transformations U_φ for $\varphi = \arccos(\frac{\cos\xi\cos\eta}{\sqrt{\cos^2\xi\cos^2\eta+\sin^2\xi\sin^2\eta}})$. Except the cases when $\eta = k\pi/2$ (for $k = 0, 1, 2, 3, \ldots$), each unitary operation U_φ is realized. Therefore, either I, or σ_z is realized on measurement-assisted CNOT processor with the smallest probability depending on whether $\sin^2\eta$ is larger than $\cos^2\eta$, or not. In particular $P^{{\rm CNOT},\eta}_{\rm success} = \min\{\cos^2\eta, \sin^2\eta\}$. The average success probabilities for measurement-assisted CNOT are given as follows

$$\overline{P}^{{\rm CNOT},\eta}_{\rm success} = \frac{1}{2\pi}\left(\int_0^{2\pi} {\rm d}\xi \sin^2\xi + \cos^2\eta \int_0^{2\pi} {\rm d}\xi \cos 2\xi\right) = \frac{1}{2}\,.$$

A straightforward calculation shows that in the *average* sense all measurement-assisted CNOT processors are equivalent. On the other hand they are not equivalent in the sense of the worst success probabilities.

It is not easy to calculate these quantities for the case when we are allowed to alternate (optimize) measurements in order to increase the success probability. In particular, it is difficult to express analytically the dependence $\xi = \xi(\varphi)$ and consequently to find an analytic expression for the optimal success probability $P_{\rm success}(U_\varphi)$. It is clear that each of the transformations can be realized with the probability strictly larger than 1/2, i.e. $P^{\rm CNOT}_{\rm success} > 1/2$ and also $\overline{P}^{\rm CNOT}_{\rm success} > 1/2$.

Let us assume that we want to perform unitary transformations U_φ approximatively. What is then the optimal program state ξ_φ approximating the given U_φ on the CNOT processor? Denoting by $\mathcal{E}_p[\varrho] = p\varrho + (1-p)\sigma_z\varrho\sigma_z$ and using the derived expression for the process fidelity between an arbitrary channel and a unitary transformation we get for the error

$$\epsilon(U_\varphi) \;=\; 1 - \max_p F(U_\varphi, \mathcal{E}_p) = 1 - \frac{1}{4}\max\{|{\rm Tr}U^\dagger_\varphi|^2, |{\rm Tr}U^\dagger_\varphi\sigma_z|^2\}\,, \tag{3.1}$$

i.e. it is optimal to encode the operation U_φ into one of two program states encoding the identity and σ_z operation, i.e. state $|\pm\rangle$. The encoding is optimal for the state for which the overlap between the desired transformation and the program I (or σ_z) is larger. In particular, $\epsilon(U_\varphi) = 1 - \max\{\cos^2\varphi, \sin^2\varphi\}$. That is, the characteristics of the approximate programmable CNOT processor describing the quality of the realization of the set of unitary transformations U_φ are the following

$$\epsilon_{\rm CNOT} \;=\; \max_\varphi \epsilon(U_\varphi) = 1/2\,, \tag{3.2}$$

$$\overline{\epsilon}_{\rm CNOT} \;=\; \frac{1}{2\pi}\int {\rm d}\varphi\,\epsilon(U_\varphi) = 1/2 - 1/\pi\,. \tag{3.3}$$

One can directly verify that the bound $P_{\rm error} \geq \epsilon$ holds in all its variations. It is saturated only for the worst case typical for the optimal measurement-assisted CNOT processor compared with the worst case approximation error, when we have $P^{{\rm CNOT},\eta=\pi/4}_{\rm error} = \epsilon_{\rm CNOT} = 1/2$.

B. Quantum information distributor

Originally, the quantum information distributor (QID) was introduced as a device performing the optimal quantum cloning and the optimal quantum NOT operation. In [11] the authors showed that the QID can be used as a universal measurement-assisted processor realizing all unitary transformations. The QID is a device with a data register representing by a qudit (d dimensional quantum system) and program register represented by two qudits, i.e. $N = d^2$. It belongs to the family of U processors [3], or equivalently the controlled-U gates. It is defined as $G_{\rm QID} = \sum_k U_k \otimes |\theta_k\rangle\langle\theta_k|$, where $k = 0, \ldots, d^2-1$, $U_k \equiv U_{ab} = \sum_r e^{i2\pi ar/d}|r-b\rangle\langle r|$ (for qubits sigma operators), $|\theta_k\rangle = U_k \otimes I|\theta_0\rangle$, $|\theta_0\rangle = \frac{1}{\sqrt{d}}\sum_j |j\rangle \otimes |j\rangle$. In other words, the unitary operators U_k form an orthogonal operator basis ${\rm Tr}U^\dagger_j U_k = d\delta_{jk}I$, and the states $|\theta_k\rangle$ form the basis of two-qudit Hilbert space composed of maximally entangled states.

In particular, for qubit $U_k = \sigma_k$ and $\{|\theta_k\rangle\}$ is the Bell basis. The program states $|\theta_k\rangle$ encode unitary transformations U_k. A general program state ξ induces one of the generalized Pauli channels $\mathcal{E}_\xi[\varrho] = \sum_k p_k U_k \varrho U_k^\dagger$ with $p_k = \langle\theta_k|\xi|\theta_k\rangle$.

Using the measurement basis $|m_j\rangle \equiv |m_{xy}\rangle = |-x\rangle \otimes \frac{1}{\sqrt{d}}\sum_r e^{i2\pi y/d}|r-x\rangle$ one can probabilistically implement an arbitrary unitary transformation with the probability $P_{\rm success} = 1/d^2$ [12]. The result $|m_{00}\rangle = \frac{1}{d^2}\sum_k |\theta_k\rangle = |0\rangle \otimes \frac{1}{\sqrt{d}}\sum_r |r\rangle$ indicates the successful realization of the program U. In particular, using this measurement basis the action of the processor can be written in the following form

$$G_{\rm QID}|\psi\rangle \otimes |\Xi\rangle = \frac{1}{d^2}\sum_j U_j A(\Xi) U_j^\dagger |\psi\rangle \otimes |m_j\rangle\,, \tag{3.4}$$

where $A(\Xi) = \sum_k \langle\theta_k|\Xi\rangle U_k$. If $A(\Xi)$ is a unitary operation then probabilities of outcomes m_j are data independent and they read $p(m_j) = \frac{1}{d^2}$. The described measurement-assisted quantum processor implements unitaries with the success probability

$$\overline{P}^{{\rm QID},M}_{\rm success} = P^{{\rm QID},M}_{\rm success} = 1/d^2\,. \tag{3.5}$$

It is of interest to show whether this is the optimal measurement-assisted probabilistic realization, or not. Note that this processor is indeed very specific, because it performs all unitaries with the same probability $P_{\rm success}(U) = 1/d^2$. A similar conclusions like in the case of CNOT processor can be made. In particular, this measurement-assisted QID processor is optimal in the sense of worst case optimality, because using different measurements will result in nontrivial success probability distributions over the set of unitary transformations.

The approximate realization is again similar to the case of CNOT gate. Both of them belong to the family of U processors, i.e. they are of the form $G = \sum_j U_j \otimes |j\rangle\langle j|$, where U_j are arbitrary unitary operations and program states $\{|j\rangle\}$ form an orthonormal basis of $\mathcal{H}_p$. These processors are not very "rich" from the point of view of approximative implementation of all unitary transformations. In fact, these processors deterministically perform the random unitary channels $\mathcal{E}_{\vec{p}}[\varrho] = \sum_j p_j U_j \varrho U_j^\dagger$ with $p_j = \langle j|\xi|j\rangle$ providing that the program register was initially prepared in the state ξ. As we shall see for the approximative programming only the basis states $|j\rangle$ are useful. In fact

$$\epsilon(U) = \max_{\vec{p}} F(U, \mathcal{E}_{\vec{p}}) = 1 - \frac{1}{d^2}\sum_j p_j |{\rm Tr} U^\dagger U_j|^2 = 1 - \frac{1}{d^2}\max_j\{|{\rm Tr}U^\dagger U_j|^2\}\,. \tag{3.6}$$

For a unitary U the best approximation is achieved by the program state $|j\rangle$ maximizing the overlap $|{\rm Tr}U^\dagger U_j|^2$. For the QID processor an arbitrary U can be written as a linear combination of the operators U_k, i.e. $U = \sum_k \alpha_k U_k$ and $\epsilon(U) = 1 - \frac{1}{d^2}\max_j |\alpha_k {\rm Tr} U_k U_j|^2 = 1 - \max_j |\alpha_j|^2$. Thus, the quality of approximations on the QID processor is given by

$$\epsilon_{\rm QID} = \max_U \epsilon(U) = 1 - 1/d^2\,. \tag{3.7}$$

This value is achieved for the unitary transformation $U = \frac{1}{2}\sum_j U_j$.

C. SWAP

The quantum SWAP gate acts on two registers of the same size, $d = N$, and its performance is defined in the following way $G_{\rm SWAP}(\varrho\otimes\xi)G^\dagger_{\rm SWAP} = \xi\otimes\varrho$. Taking the SWAP gate as a processor we can implement all contractions to a single point specified by the program state, i.e. $\mathcal{C}_\xi[\varrho] = \xi$ for all input states ϱ. It is clear that such processor does not belong to the family of U processors. Indeed it cannot be used to realize any unitary transformation. However, unlike for U processors, for each program state the SWAP processor implements different program, which makes it exceptional, because there is no redundancy in the state space of the program register. Let remind us that for U processors all states having the same diagonal elements in the basis of vectors $|j\rangle$ encode the same quantum operation. The set of contractions is closed under convex combinations and except extremal points of the whole set of quantum operations (pure-state contractions) they contain also totally random channel, i.e. a contraction to the total mixture.

Let us consider an arbitrary measurement M of the program register. The probability of measuring the result $|m\rangle$ is given as

$$p(m) = {\rm Tr}[Q_m G_{\rm SWAP}(\varrho\otimes\xi)G^\dagger_{\rm SWAP} Q_m] = {\rm Tr}[\xi\otimes|m\rangle\langle m|\varrho|m\rangle\langle m|] = \langle m|\varrho|m\rangle\,,$$

where we used the notation $Q_m = I \otimes |m\rangle\langle m|$. That is, for any measurement the resulting probability distribution depends on the initial state of the data, i.e. no specific quantum operation is associated with particular results. Or, to be more precise, for each outcome the contraction to a fixed point ξ is realized. Therefore, measurement-assisted SWAP processors are in some sense trivial and do not provide us with any improvement of performance. Therefore , it does not make much sense to apply measurements at the output of the SWAP processor. Consequently, the SWAP processor is not suitable for performing probabilistically any unitary transformation, i.e. $P_{\rm success}^{\rm SWAP} = \overline{P}_{\rm success}^{\rm SWAP} = 0$.

In what follows we will see that from the point of view of approximative implementation of unitary transformations, the SWAP processor acts much better. Let us calculate the error $\epsilon(U) = 1 - \frac{1}{d^2}\max_\Xi \sum_r |{\rm Tr}U^\dagger A_r(\Xi)|^2$, where $A_r(\Xi)$ are Kraus operators associated with pure program states, i.e. $A_r(\Xi) = \langle r|G_{\rm SWAP}|\Xi\rangle = \langle r|\sum_{j,k}|k\rangle\langle j| \otimes |j\rangle\langle k|\Xi\rangle = |\Xi\rangle\langle r|$. In the matrix (basis) representation of $G_{\rm SWAP}$ one can use any basis, i.e. even a basis containing the state vector $|\Xi\rangle$. Our task is to find $\epsilon(U) = 1 - \frac{1}{d^2}\max_\Xi \sum_r |\langle r|U^\dagger|\Xi\rangle|^2$. Choosing the basis containing the vector $|\Xi\rangle$ we obtain that we have to maximize the length of the column of the unitary transformation $U^\dagger$. But we know that columns form mutually orthonormal vectors, and therefore in this basis the expression $\sum_r |\langle r|U^\dagger|\Xi\rangle|^2 = 1$. This means that each pure program state $|\Xi\rangle$ approximates each unitary operation with the same accuracy measured by $\epsilon(U) = 1 - 1/d^2$. And consequently

$$\epsilon_{\rm SWAP} = \overline{\epsilon}_{\rm SWAP} = 1 - 1/d^2 \,. \tag{3.8}$$

Thus, the worst case accuracy of the QID and the SWAP processors are the same. Nevertheless, it should be noted that for the QID the program space is twice as large as the program space for the SWAP processor. Therefore, we can conclude that the SWAP processor is more optimal (suitable) for approximate programming. However, the situation is completely different for probabilistic programming.

IV. OPTIMALITY OF APPROXIMATE PROCESSORS

Each quantum processor induces a mapping $\mathcal{G}$ from the set of program states into a subset of all quantum operations applied on the data register, i.e. $\mathcal{G} : \xi \mapsto \mathcal{E}_\xi$. Let us denote by $\Gamma_\mathcal{G} \subset \Gamma$ the subset of deterministically implementable quantum programs, where Γ stands for the set of all possible quantum operations. Since $\Gamma_\mathcal{G}$ is a linear image of the set of program states, it is convex. The question of optimality then can be illustrated in the following way. Denote by $\partial\Gamma$ the boundary with respect to some topology. Then the worst case optimality parameter measures the distance between the points of the sets $\Gamma_\mathcal{G}$ and $\partial\Gamma$. Formally, $\epsilon_G = \max_{\mathcal{T}\in\Gamma_\mathcal{G}} \min_{\mathcal{E}\in\Gamma} D(\mathcal{E},\mathcal{T})$.

For the process fidelity we have

$$\epsilon_G = 1 - \min_{\mathcal{T}\in\Gamma_\mathcal{G}} \max_{\mathcal{E}\in\Gamma} F(\mathcal{E},\mathcal{T}) \,. \tag{4.1}$$

Let us consider $\mathcal{E}', \mathcal{T}'$ that optimize the accuracy. Because of the fact that $F(\sum_k p_k\mathcal{E}_k, \mathcal{T}') \geq \sum_k p_k F(\mathcal{E}_k, \mathcal{T}')$ it follows that $\mathcal{E}'$ can be always chosen to be the extremal quantum operation of the set Γ. Let us formulate the main problem: given a data register of size d and given the program register of size N. What is the optimal approximate processor? This problem is indeed difficult and only partial results are known.

Let us pay attention to an optimal realization of all unitary transformations. In this case the approximate processor is *universal* if it performs an arbitrary unitary transformation with nonzero accuracy, i.e. $0 < \epsilon(U) \leq 1$. For instance, the CNOT processor is not universal in this sense, but the SWAP processor is. The necessary and sufficient condition for a processing being universal is the following: The process fidelity has to nonzero for all unitary transformations which means that $F(U,\xi) = \frac{1}{d^2}\sum_r |{\rm Tr}U^\dagger A_r(\xi)|^2 > 0$. The general processor can be written in the form $G = \sum_{jk} A_{jk} \otimes |j\rangle\langle k|$. Providing that A_{jk} form a complete operator basis the process fidelity is different from zero for all unitary transformations. In particular, for the SWAP processor the operators $A_{jk} = |k\rangle\langle j|$ undoubtedly form a complete basis.

For the d-dimensional data register the processor can be universal (implementing all unitaries) only if the operators A_{jk} form an operator basis, i.e. they are independent and the total number of them is d^2. This necessarily means that universal processors for the d dimensional program registers must use at least d dimensional program register. It is easy to see that U processors with such program size cannot be universal, because they do not contain sufficient amount of independent operators, $A_{jk} = \delta_{jk}U_j$. The existence of an approximate universal processor of this size is guaranteed by an example of the SWAP processor. The question is whether this processor is optimal, i.e. whether $\epsilon_{\rm SWAP} = 1-1/d^2$ attains indeed the minimal value for a universal processor of the program size $N = d$. D'Arianno and Perrinoti [6, 14] found that this is indeed the case for a single qubit (although they considered a different processor).

In some sense such an error can be achieved trivially. The process fidelity will be nonzero for all quantum operations providing that at least for one of the program states ξ encodes an operation $\mathcal{T}$ such that the state $\Phi_\mathcal{T}$ has the full rank, i.e. $\mathrm{Tr}\sqrt{\sqrt{\Phi_{\mathcal{E}}}\Phi_{\mathcal{T}}\sqrt{\Phi_{\mathcal{E}}}} > 0$ for all $\mathcal{E}$. This condition is only sufficient, but not necessary for a processor to be universal. Anyway, once we are discussing the problem in general, this condition is sufficient to derive the the lower bound on the accuracy of approximate processors. Consider that the processor enables us to perform a contraction into the total mixture, i.e. let us consider the map $\mathcal{A}[\varrho] = \frac{1}{d}I$. Then for any $\mathcal{E}$ the process fidelity reads $F(\mathcal{A},\mathcal{E}) = \frac{1}{d^2}[\mathrm{Tr}\sqrt{\Phi_{\mathcal{E}}}]^2$, because $\Phi_{\mathcal{A}} = \frac{1}{d^2}I$. The minimum is achieved for pure states $\Phi_{\mathcal{E}}$ corresponding to unitary transformations. That is, all other programs are approximated by this operation better than unitaries and therefore for such universal processor $\epsilon_G \geq 1 - 1/d^2$. The SWAP as well as the QID processors can realize the transformation $\mathcal{A}$ and they both saturate this bound.

Let us restrict the original problem and ask the following question: What is the optimal U processor? To be able to discuss optimality processors they have to be universal. It means that this question does not make any sense for program registers of the size $N < d^2$. The QID is an example of an approximate processor in the case when $N = d^2$. For the U processors the situation is simpler because the process fidelity is given by the formula $F(U,\mathcal{E}) = \frac{1}{d^2}\sum_r |\mathrm{Tr}U^\dagger A_r|^2$. Let us fix the dimension of the program space to be $N = d^2$, i.e. we work with d^2 unitaries U_j. The first question is which of the U processors is optimal in implementing all unitary transformations. We know that the best approximation is achieved for the basis states $|j\rangle$ and $\epsilon(U) = 1 - \frac{1}{d^2}\max_j\{|\mathrm{Tr}U^\dagger U_j|^2\}$. That is, the worst case is represented by a unitary U_x having the smallest maximal overlap with "elementary" programs U_j. Our aim is to show that this worst case is optimized for mutually orthogonal operators σ_j, i.e. $\mathrm{Tr}\sigma_j^\dagger\sigma_k = d\delta_{jk}$. Processors having as elementary programs mutually orthogonal unitaries are all equivalent. In particular, the approximation of the operator $U_x = \frac{1}{d}\sum_j \sigma_j$ maximize the error to $\epsilon(U_x) = 1 - 1/d^2$. Each unitary transformation can be expressed in a suitable orthogonal operator basis in such form, i.e. each unitary operation can represent the worst case for some processor.

Consider that we know the transformation U_x for a processor with nonorthogonal unitaries. Let us define a processor using orthogonal operators σ_j such that $U_x = \frac{1}{d}\sum_j \sigma_j$. Let us express also the elementary operators U_j in this new operator basis $U_j = \sum_a u_{ja}\sigma_a$. Calculating the overlap we obtain the following bound

$$|\mathrm{Tr}U_x U_k|^2 = |\frac{1}{d}\sum_j u_{kj}|^2 \leq \sum_j |u_{kj}|^2 \leq 1\,, \tag{4.2}$$

where we used the identities $\mathrm{Tr}\sigma_a^\dagger\sigma_b = d\delta_{ab}$ and $\mathrm{Tr}U_k^\dagger U_k = d\sum_j |u_{kj}|^2 = d$. Consequently, the error is minimized for $|\mathrm{Tr}U_x\sigma_k|^2 = 1$ and therefore we can conclude that the U processors with orthogonal elementary programs are indeed optimal, i.e. the QID processor is an optimal approximate U-processor for $N = d^2$ dimensional program register.

This optimality result for approximate processors can be used directly for probabilistic processors by using the inequality $P_{\mathrm{error}} \geq \epsilon$, or equivalently $P_{\mathrm{success}} \leq 1 - \epsilon$. The problem is now the following: Which of the U processors is optimal in a probabilistic realization of unitary transformations? Based on the answer to the similar question for approximative programming we can say that the optimal success probability is less or equal to $1 - \epsilon = 1/d^2$, i.e. $P_{\mathrm{success}} \leq 1/d^2$. However, we know that the measurement-assisted QID processor saturates this bound, i.e. $P_{\mathrm{success}}^{\mathrm{QID},M} = 1/d^2$, and therefore the QID is an example of the optimal probabilistic U processor implementing all unitary transformations.

V. OPTIMALITY FOR U(1) ROTATIONS

In this section we will relax the universality condition and we will study an implementation of a one-parametric group of unitary transformations. An example of the universal processor performing this task is the CNOT processor. Our aim will be to specify the dependence of the approximation error and the success probability on the size N of the program register.

Let us start with the approximative realization of $U_\varphi = \exp(i\varphi A)$ on U processors, i.e. we use processors of the form $G = \sum_j U_j \otimes |j\rangle\langle j|$, where $j = 1,\ldots,N$. Since the approximation error is specified by overlaps between U_φ and U_j, it is reasonable to assume that U_j are from the linear span of unitaries U_φ. Otherwise, we would obtain $\mathrm{Tr}U_\varphi^\dagger U_j = 0$ which is not interesting for approximative realization of U_φ, i.e. the associated states $|j\rangle$ do not approximate any transformation U_φ.

In particular, U_j are from the group U_φ. The overlap between two unitaries $U_\varphi, U_{\varphi'}$ is given by the expression $|\mathrm{Tr}U_\varphi^\dagger U_{\varphi'}|^2 = |\mathrm{Tr}\exp[i(\varphi' - \varphi)A]|^2 = |\sum_n e^{i(\varphi'-\varphi)a_n}|^2$, where a_n are eigenvalues of A. The closer the angles φ', φ are, the larger is the overlap. In what follows we will assume that the data register is two-dimensional, i.e. $A = \vec{a}\cdot\vec{\sigma}$, $U_\varphi = \cos\varphi I + i\sin\varphi\,(\vec{a}\cdot\vec{\sigma})$ and $|\mathrm{Tr}U_\varphi^\dagger U_{\varphi'}|^2 = 4\cos^2(\varphi' - \varphi)$.

Let us consider a processor with N dimensional program register performing N unitaries U_{φ_j}. The question is, what are the best choices of angles φ_j in order to minimize the approximation error of the implementation of the whole group U_φ ($\varphi \in [0, 2\pi]$). The program state $|j\rangle$ best approximates those angles φ that lie in the vicinity of the angle φ_j (the difference $\varphi_j - \varphi$ is smaller than the differences $\varphi_k - \varphi$ for $k \neq j$). Because of the identity $\cos^2(\varphi_j - \varphi + \pi) = \cos^2(\varphi_j - \varphi)$ the state $|j\rangle$ approximates with the same accuracy the unitaries U_φ and $U_{\varphi+\pi}$. That is, it is enough to consider only implementation of unitaries specified by angles within the interval $\varphi \in [0, \pi]$. The whole problem can be illustrated in the following way. The angles φ_j divide the half-circle (angles from 0 to π) into N regions. The state $|j\rangle$ approximates optimally all the angles from the interval $[\varphi_j - \frac{1}{2}(\varphi_j - \varphi_{j-1}), \varphi_j + \frac{1}{2}(\varphi_{j+1} - \varphi_j)]$ (for simplicity we assume that $\varphi_1 = 0$ and $\varphi_j < \varphi_{j+1}$). The best choice of φ_j is such that the half-circle is divided into equally large regions, i.e. the angles φ_j are separated by the same angle $\varphi_{j+1} - \varphi_j = \pi/N$. In such case the approximation error reads

$$\epsilon_G = 1 - \frac{1}{4} 4 \cos^2(\pi/(2N)) = 1 - \cos^2(\pi/(2N)) \tag{5.1}$$

because for arbitrary φ the smallest difference is $\min_j\{\varphi_j - \varphi\} \leq \pi/(2N)$, i.e. the overlap is maximal and the error ϵ_G is optimal. The average error of implementation of the operations U_φ equals to

$$\overline{\epsilon}_G = \frac{1}{2}(1 - \frac{N}{\pi} \sin \frac{\pi}{N}) \,. \tag{5.2}$$

Based on this result we can bound the optimal probabilistic realization of qubit operations U_φ using the N dimensional program register and the U processor

$$P^G_{\rm success} \; \leq \; \cos^2(\frac{\pi}{2N}) \,, \tag{5.3}$$
$$\overline{P}^G_{\rm success} \; \leq \; \frac{1}{2}(1 + \frac{N}{\pi} \sin \frac{\pi}{N}) \,. \tag{5.4}$$

Now the question is whether there exists a measurement-assisted U processors saturating these bounds. In what follows we will present an example of the U processor that saturates this bound in a limit sense, i.e. for large program registers.

In Refs. [15, 16] the authors proposed a way how to increase the success probability to arbitrarily close to unity by using large program registers with the dimensionality $N = \dim \mathcal{H}_{\rm p} = 2^n$, where n is the number of qubits in the program register. They found a network of the controlled U form composed of n sequentially applied k-Toffoli gates for $k = 1, 2 \ldots n$. The k-Toffoli (T_k) implements a controlled NOT operation with k control qubits and a single target qubit. The whole processor acts on $n + 1$ qubits, where the first qubit represents the data and all other qubits form the program register. The k-Toffoli gate uses qubits $1, \ldots k$ as control qubits (this includes the program qubit as the qubit number 1) and the $(k + 1)$th qubit is the target qubit. The whole processor is described by the operator $G = T_n(T_{n-1} \otimes I) \ldots (T_2 \otimes I^{\otimes(n-2)})(T_1 \otimes I^{\otimes(n-1)})$. Vidal et al. showed [15, 16] that it is possible to choose program states and a measurement such that the success probability scales as

$$P^M_{\rm success}(U_\varphi) = 1 - (1/N) \tag{5.5}$$

for an arbitrary operation $U_\varphi = e^{i\varphi\sigma_z}$. Thus, we find that $P^{G,M}_{\rm success} = \overline{P}^{G,M}_{\rm success} = 1 - (1/N)$, which is consistent with the derived bound, but scales differently, because $P^G_{success} \leq 1 - \sin^2(\pi/2N) \to 1 - (\pi/(2N))^2$ for large N. This means that the saturation of the bound is still an open question. Moreover, let remind us that the derived formula for the measurement-assisted U processor holds only for dimensions $N = 2^n$, where n is the size of the program register in the number of qubits.

VI. CONCLUSION

One of the goals of the research in the field of quantum computing is a construction of programmable quantum processor. It follows from the work of Nielsen and Chuang that perfect programmable processor does not exist. In this paper we have analyzed two different scenarios relaxing the condition of universality: approximate implementation of quantum programs (approximate processors) and probabilistic implementation of quantum programs (probabilistic processors). We have discussed the problems of optimality of universal approximate processors and universal (measurement-assisted) probabilistic processors. Using examples of the CNOT, the QID and the SWAP processors we have shown explicitly the validity of the general relation between the approximation error ϵ_G and the success probability $P^G_{\rm success}$, respectively. In particular, we have seen that the inequality $P_{\rm error} \geq \epsilon$ is not saturated in general case. For instance, the SWAP processor is completely useless in probabilistic programming $P^{\rm SWAP}_{\rm error} = 1$, but approximatively it performs optimally.

We have studied in detail optimality of restricted class of processors - the so called U processors. We have shown that the QID processor is optimal in implementing unitary transformations on a qudit with $N = d^2$ dimensional program register. Under the same settings we have analyzed the restricted universality of an implementation of only one parametric set of unitary transformations. Unlike the QID processor we have found that the presented example of the U processor does not saturate the bound derived from approximative implementation. That is, the question of optimality for this processor is an open question.

The optimality questions of universal processors (either approximate, or probabilistic) represent a difficult open problem of quantum information science [18]. In this paper, we have described current state of the art and we have presented known results. The question of efficient programmability of quantum computers makes these problems of optimality very attractive and they deserve further investigation.

Acknowledgements

This work was supported in part by the European Union projects INTAS-04-77-7289 and QAP, by the Slovak Academy of Sciences via the project CE-PI I/2/2005 and VEGA 2/6070/26, by the project APVT-99-012304 and GAČR GA 201/01/0413.

[1] J. Gruska, *Foundations of Computing* (Thompson Computer Press, London, 1999).
[2] M. Nielsen and I. Chuang, Phys. Rev. Lett. **79**, 321 (1997).
[3] M. Hillery, M. Ziman, and V. Bužek, Phys. Rev. A **66**, 042302 (2002).
[4] A. Yu. Vlasov, quant-ph/0103119; quant-ph/0311196; quant-ph/0301147; quant-ph/0503230.
[5] M. Dušek and V. Bužek, Phys. Rev. A **66**, 022112 (2002).
[6] G. M. D' Ariano and P. Perinotti, Phys. Rev. Lett. **94**, 090401 (2005).
[7] J. Gea-Banacloche, Phys. Rev. A bf 65, 022308 (2002).
[8] A. Silberfarb and I. Deutsch, Phys. Rev. A **68**, 013817 (2003).
[9] A. K. Ekert, C. M. Alves, D. K. L. Oi, M. Horodecki, P. Horodecki, and L. C. Kwek, Phys. Rev. Lett. **88**, 217901 (2002).
[10] J. P. Paz and A. Roncaglia, Phys. Rev. A **68**, 052316 (2003).
[11] M. Hillery, M. Ziman, and V. Bužek, Phys. Rev. A **65**, 022301 (2002).
[12] M. Hillery, M. Ziman, and V. Bužek, Phys. Rev. A **69**, 042311 (2004).
[13] M. Hillery, M. Ziman, and V. Bužek, Phys. Rev. A **73**, 022345 (2006).
[14] G. M. D' Ariano and P. Perinotti, quant-ph/0510033.
[15] G. Vidal and J. I. Cirac, quant-ph/0012067.
[16] G. Vidal, L. Masanes, and J. I. Cirac, Phys. Rev. Lett. **88**, 047905 (2002).
[17] M. Ziman and V. Bužek, Int. J.Quant. Inf. **1**, 523 (2003)
[18] M.A. Nielsen and I.L. Chuang, *Quantum Computation and Quantum Information*, (Cambridge University Press, Cambridge, 2000)