PHYSICAL REVIEW A **95**, 062325 (2017)

Robust quantum gates for stochastic time-varying noise

Chia-Hsien Huang and Hsi-Sheng Goan*
Department of Physics and Center for Theoretical Sciences, National Taiwan University, Taipei 10617, Taiwan
and Center for Quantum Science and Engineering, National Taiwan University, Taipei 10617, Taiwan

(Received 4 October 2016; published 16 June 2017)

How to effectively construct robust quantum gates for time-varying noise is a very important but still outstanding problem. Here we develop a systematic method to find pulses for quantum gate operations robust against both low- and high-frequency (comparable to the qubit transition frequency) stochastic time-varying noise. Our approach, taking into account the noise properties of quantum computing systems, can output single smooth pulses in the presence of multisources of noise. Furthermore, our method can be applied to different system models and noise models and will make essential steps toward constructing high-fidelity and robust quantum gates for fault-tolerant quantum computation. Finally, we discuss and compare the gate operation performance by our method with that by the filter-transfer-function method.

DOI: 10.1103/PhysRevA.95.062325

I. INTRODUCTION

To realize practical quantum computation, a set of high-fidelity universal quantum gates robust against noise in the qubit system is prerequisite. Constructing control pulses to operate quantum gates which meet this requirement is an important and timely issue. Quantum gates in open quantum systems have been investigated by various methods such as dynamical decoupling methods [1–6] and optimal control methods [7–10]. For classical noise, there are many robust control methods such as composite pulses [11–19], soft uniaxial positive control for orthogonal drift error (SUPCODE) [20–25], the sampling-based learning control method [26–28], inhomogeneous control methods [29,30], the analytical method [31], the single-shot pulse method [32], optimal control methods [33–35], the invariant-based inverse engineering method [36,37], and filter-transfer-function (FTF) methods [38–41]. However, in most of these methods [11–34], noise is assumed to be quasistatic, i.e., is time independent within the gate operation time but can vary between gates. We call these robust control strategies quasistatic-noise (QSN) methods. But this QSN assumption is not always valid [42]. The robust performance of control pulses obtained by QSN methods under time-dependent noise (e.g., $1/f^\alpha$ noise) [22,23,25,43] has been investigated, and it was found that they can still work well for relatively low-frequency non-Markovian noise (e.g., $1/f^\alpha$ noise with $\alpha \gtrsim 1$).

Stochastic time-dependent noise is treated with the FTF method [38–40], in which the area of the filter-transfer function in the frequency region, where the noise power spectral density (PSD) is nonnegligible, is minimized. However, in this approach only the the filter-transfer function overlapping with the noise PSD in the preset frequency region is considered, and the detailed information on the distribution of the noise PSD is not included in the optimization cost function. Here we develop an optimal control method in the time domain by choosing the ensemble average gate infidelity (error) as our cost function for optimization. As a result, the noise correlation function (CF), or, equivalently, the detailed noise PSD distribution, appears naturally in our chosen optimization cost function. Therefore our method can show a more robust performance against noise in the general case. The idea of our method is simple, and our method is not limited to particular system models, noise models, or noise CFs. We demonstrate our robust control method for classical noise in this paper, but our method can be easily generalized to the case of quantum noise by replacing the ensemble average for classical noise with the trace over the degrees of freedom of the quantum noise (environment) [41]. In other words, our method can be applied to systems with both classical noise and quantum noise present simultaneously.

II. ENSEMBLE AVERAGE INFIDELITY AND OPTIMIZATION METHOD

We first introduce our robust control method here and then compare it with the QSN method and the FTF method. We consider a total Hamiltonian $\mathcal{H}(t) = \mathcal{H}_I(t) + \mathcal{H}_N(t)$, where $\mathcal{H}_I(t)$ is the ideal system Hamiltonian and $\mathcal{H}_N(t)$ is the noise Hamiltonian. If a system is ideal, i.e., $\mathcal{H}_N(t) = 0$, then its propagator is $U_I(t) = \mathcal{T}_+ \exp[-i\int_0^t \mathcal{H}_I(t')dt']$ (throughout this paper we set $\hbar = 1$), where $\mathcal{T}_+$ is the time-ordering operator. However, in reality, there may be many sources of noise present in the system, so $\mathcal{H}_N(t) = \sum_j \beta_j(t) H_{N_j}(t)$, where $\beta_j(t)$ is the strength of the jth stochastic time-varying noise (TVN) and $H_{N_j}(t)$ is the corresponding system coupling operator term. The propagator for a realistic system is then $U(t) = U_I(t)\mathcal{T}_+ \exp[-i\int_0^t \tilde{\mathcal{H}}_N(t')dt']$. Here $\tilde{\mathcal{H}}_N(t) = \Sigma_j \beta_j(t) R_j(t)$ is the noise Hamiltonian in the interaction picture transformed by $U_I(t)$ and $R_j(t) \equiv U_I^\dagger(t) H_{N_j}(t) U_I(t)$. Suppose that U_T is our target gate and the gate operation time is t_f. The gate infidelity (error) $\mathcal{I}$ for an n-qubit gate can be defined as

$$\mathcal{I} \equiv 1 - \frac{1}{4^n}|\mathrm{Tr}[U_T^\dagger U(t_f)]|^2, \tag{1}$$

where Tr denotes a trace over the n-qubit system state space. If the noise strength is not too strong, we can expand the propagator $U(t_f)$ in terms of $\tilde{\mathcal{H}}_N(t)$ by Dyson series [44] into the form $U(t_f) = U_I(t_f)[I + \Psi_1 + \Psi_2 + \cdots]$, where the first two terms of Ψ_j are $\Psi_1 = -i\int_0^{t_f} \tilde{\mathcal{H}}_N(t')dt'$ and $\Psi_2 = -\int_0^{t_f} dt_1 \int_0^{t_1} dt_2 \tilde{\mathcal{H}}_N(t_1)\tilde{\mathcal{H}}_N(t_2)$. Substituting the expanded $U(t_f)$ into $\mathcal{I}$ in Eq. (1), the expanded infidelity $\mathcal{I}$

---

*goan@phys.ntu.edu.tw

(see Appendix A) takes the form

$$\mathcal{I} = J_1 + J_2 + \epsilon + O\big(\tilde{\mathcal{H}}_N^m, m \geqslant 3\big), \tag{2}$$

$$J_1 \equiv 1 - \frac{1}{4^n}|\mathrm{Tr}[U_T^\dagger U_I(t_f)]|^2, \tag{3}$$

$$J_2 \equiv -\frac{1}{2^{n-1}}\mathrm{Re}[\mathrm{Tr}(\Psi_2)] - \frac{1}{4^n}|\mathrm{Tr}(\Psi_1)|^2. \tag{4}$$

Here J_1 is the definition of gate infidelity for the ideal system, J_2 is the lowest-order contribution of the noise to the gate infidelity, ϵ (detailed form shown in Appendix A) denotes an extra contribution that is correlated to J_1 and the Dyson expansion terms Ψ_j, and $O(\tilde{\mathcal{H}}_N^m, m \geqslant 3)$ represents other higher-order terms excluding ϵ. If the noise strength is not too strong such that $|\Psi_{j+1}| \ll |\Psi_j|$, the extra contribution ϵ will become negligible when J_1 is getting small (see discussion in Appendix A). The symbol Re in Eq. (4) denotes taking the real part of the quantity it acts on. Because noise $\beta_j(t)$ is stochastic, we denote the ensemble average of the infidelity over the different noise realizations as

$$\langle \mathcal{I} \rangle = J_1 + \langle J_2 \rangle + \langle \epsilon \rangle + \big\langle O\big(\tilde{\mathcal{H}}_N^m, m \geqslant 3\big)\big\rangle. \tag{5}$$

Here

$$\begin{aligned} \langle J_2 \rangle = & \sum_{j,k} \int_0^{t_f} dt_1 \int_0^{t_1} dt_2 C_{jk}(t_1,t_2) \frac{\mathrm{Tr}[R_j(t_1)R_k(t_2)]}{2^{n-1}} \\ & - \sum_{j,k} \int_0^{t_f} dt_1 \int_0^{t_f} dt_2 C_{jk}(t_1,t_2) \frac{\mathrm{Tr}[R_j(t_1)]\mathrm{Tr}[R_k(t_2)]}{4^n}, \end{aligned} \tag{6}$$

where $C_{jk}(t_1,t_2) = \langle \beta_j(t_1)\beta_k(t_2)\rangle$ is the CF for noise $\beta_j(t_1)$ and $\beta_k(t_2)$. The first-order noise term proportional to $\mathrm{Re}[\mathrm{Tr}(\Psi_1)]$ vanishes due to the fact that $\mathrm{Tr}(\Psi_1)$ is purely imaginary rather than the assumption of $\langle \beta_j(t)\rangle = 0$ (see Appendix A). If different sources of noise are independent, $C_{jk}(t_1,t_2) = 0$ for $j \neq k$, and if the noise Hamiltonian $\mathcal{H}_N(t)$ is traceless, the second term in Eq. (6) vanishes. The ideal Hamiltonian $\mathcal{H}_I(t)$ is a function of the control field $\Omega(t)$, that is, $\mathcal{H}_I(t) = \mathcal{H}_I(\Omega(t))$, and the control field $\Omega(t)$ is chosen to be a function of a set of control parameters $[a_1, a_2, \ldots]$. Then $U_I(t)$ and each term of the ensemble average infidelity $\langle \mathcal{I} \rangle$ in Eq. (5) is also a function of the control parameter set $[a_1, a_2, \ldots]$. Our goal is to search the optimal parameter set $[a_1, a_2, \ldots]$, which minimizes the ensemble average infidelity $\langle \mathcal{I} \rangle$. If the noise strength or fluctuation is not large, then the dominant noise contribution to $\langle \mathcal{I} \rangle$ is from $\langle J_2 \rangle$ as the higher-order terms $\langle O(\tilde{H}_N^m, m \geqslant 3)\rangle$ can be neglected (see Appendix B). J_1 can generally be made sufficiently small that the extra term $\langle \epsilon \rangle$ in $\langle \mathcal{I} \rangle$ of Eq. (5) can be safely ignored. So we concentrate on the minimization of $\langle \mathcal{I} \rangle \cong J_1 + \langle J_2 \rangle$ to obtain the optimal control parameter set. We, however, use the full-order ensemble average infidelity $\langle \mathcal{I} \rangle$ (described later) to examine the performance of the optimal control parameter set found in this way.

We use two-step optimization to achieve this goal. The first step is called the J_1 optimization, in which J_1 is the cost function. The gate infidelities J_1 in an ideal unitary system with gate-operation controllability and a sufficient number of control parameters can be made as low as one wishes, limited only by the machine precision of the computation. So using an ensemble of random control parameter sets as initial guesses, we obtain after the J_1 optimization an ensemble of optimized control parameter sets, all with very low values of J_1. The second step is called the $J_1 + \langle J_2 \rangle$ optimization. We take $J_1 + \langle J_2 \rangle$ as a cost function and randomly choose some optimized control parameter sets in the first optimization step as initial guesses to run the optimal control algorithm. After the $J_1 + \langle J_2 \rangle$ optimization, we obtain an ensemble of control parameter sets with low values of $J_1 + \langle J_2 \rangle$ and then choose the lowest one as the optimal control parameter set. The purpose of using the two-step optimization is to improve the optimization efficiency. If we run $J_1 + \langle J_2 \rangle$ optimization directly from an ensemble of random control parameter sets, we need more optimization iterations to achieve the goal, and the success rate is relatively low compared with the two-step optimization. Besides, the $J_1 + \langle J_2 \rangle$ optimization enables us to know separately the optimized values of J_1 and $\langle J_2 \rangle$. When $\langle J_2 \rangle$ can be minimized to a very small value as in the case of static or low-frequency noise, one has to use a small time step for simulation to make J_1 smaller than $\langle J_2 \rangle$. However, for high-frequency noise, $\langle J_2 \rangle$ is hard to be minimized to a very small value, and one can instead choose a suitable larger time step to make J_1 just one or two orders of magnitude smaller than $\langle J_2 \rangle$, substantially saving optimization time, especially for multiqubits and multiple sources of noise. We use the gradient-free and model-free Nelder-Mead (NM) algorithm [45] in both the J_1 and the $J_1 + \langle J_2 \rangle$ optimization steps. However, the NM algorithm may be stuck in local traps in the $J_1 + \langle J_2 \rangle$ parameter-space topography. To overcome this problem, we use the repeating-NM algorithm in the $J_1 + \langle J_2 \rangle$ optimization step. The control parameter set from the first $J_1 + \langle J_2 \rangle$ optimization may lie in a local trap. Therefore, we add random fluctuations to this control parameter set and try to pull it out of the trap. Then we use this shifted control parameter set as an initial guess to run the second $J_1 + \langle J_2 \rangle$ optimization. We repeat the same procedure many times until the values of $J_1 + \langle J_2 \rangle$ can no longer be improved (reduced) and then output the corresponding control parameter set. Our optimization method employing the gradient-free and model-free NM algorithm is quite general, capable of dealing with different forms or structures of the ideal system Hamiltonian $\mathcal{H}_I(t)$, control field $\Omega(t)$, noise Hamiltonian $\mathcal{H}_N(t)$, and noise CF $C_{jk}(t_1,t_2)$ for a few qubit systems.

The ensemble infidelity $\langle \mathcal{I} \rangle$ we use to show robust performance of the gate as the noise strength varies is calculated using the full evolution of the total system-noise Hamiltonian and many realizations of the noise without any other approximation. By inputting the optimal control parameter set obtained by the optimization strategy into the total system-noise Hamiltonian $\mathcal{H}(t) = \mathcal{H}_I(t) + \mathcal{H}_N(t)$ to obtain numerically the full propagator for a single noise realization, we can calculate the gate infidelity $\mathcal{I}$ using Eq. (1) for the noise realization. The procedure is repeated for many different noise realizations. Then we take an ensemble average of the infidelities over the different noise realizations to obtain $\langle \mathcal{I} \rangle$.

In principle, we could deal with any given form of the noise correlation function (or, equivalently, the noise PSD) to insert into Eq. (6) for the $J_1 + \langle J_2 \rangle$ optimization. But as a particular example, we choose the Ornstein-Uhlenbeck (OU) process

$\beta_{\mathrm{OU}}(t)$ to simulate stochastic TVN [46]. Studying the influence of and developing robust strategies against time-dependent noise is an important subject of research in quantum control problems both theoretically and experimentally [38–40,42,43]. If the initial noise $\beta_{\mathrm{OU}}(t=0)$ is a normal distribution with zero mean and with standard deviation σ_{OU}, then the noise CF of the OU process $\beta_{\mathrm{OU}}(t)$ is

$$C_{\mathrm{OU}}(t_1,t_2) = \sigma_{\mathrm{OU}}^2 \exp\left(-\gamma_{\mathrm{OU}}|t_1 - t_2|\right) \tag{7}$$

with the noise correlation time $\tau \sim (1/\gamma_{\mathrm{OU}})$, and the corresponding noise PSD is Lorentzian

$$S_{\mathrm{OU}}(\omega) = \frac{2\sigma_{\mathrm{OU}}^2\gamma_{\mathrm{OU}}}{\left(\gamma_{\mathrm{OU}}^2 + \omega^2\right)}. \tag{8}$$

Lorentzian PSDs of spin noise resulting in a fluctuating magnetic field at the location of the qubits in InGaAs semiconductor quantum dots have been measured experimentally [47,48]. Generally, a small γ_{OU} corresponds to low-frequency or quasistatic noise; a large γ_{OU} corresponds to high-frequency noise. The noise $\beta_{\mathrm{OU}}(t)$ can be simulated through the formula $\beta_{\mathrm{OU}}(t+dt) = (1-\gamma_{\mathrm{OU}}dt)\beta_{\mathrm{OU}}(t) + \sigma_{\mathrm{OU}}\sqrt{2\gamma_{\mathrm{OU}}}dW(t)$, where $W(t)$ is a Wiener process [46]. Figures 1(c), 1(d), and 1(e) show the different realizations of the noise $\beta_{\mathrm{OU}}(t)$ with $\sigma_{\mathrm{OU}} = 10^{-3}$ for different values of $\gamma_{\mathrm{OU}}/\omega_0 = 10^{-7}, 10^{-3}$, and 10^{-1}, respectively, where ω_0 is the typical system frequency. We note here that the particular choice of the OU noise should by no means diminish the value of our work or the power of our method. Any given or experimentally measured well-behaved noise PSD or noise CF can be dealt with. We demonstrate later that our method can also work effectively for another form of noise PSD different from that for OU noise when we compare the performance of our method with that of the FTF method. The reason for using the OU noise in the system-noise Hamiltonian here is that it is relatively easy to simulate its stochastic noise realizations in the time domain. Therefore, we can calculate the full-order ensemble average infidelity $\langle\mathcal{I}\rangle$ to show that our $J_1 + \langle J_2\rangle$ optimization, which minimizes the second-order noise contribution to the average infidelity $\langle\mathcal{I}\rangle$, can indeed work rather well for not too strong a noise fluctuation.

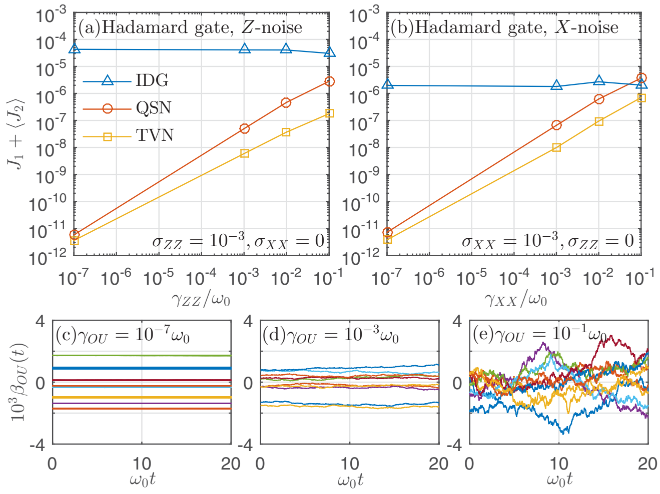

FIG. 1. $J_1 + \langle J_2\rangle$ versus (a) γ_{ZZ} for Z noise ($\sigma_{ZZ} = 10^{-3}, \sigma_{XX} = 0$) and (b) γ_{XX} for X noise ($\sigma_{XX} = 10^{-3}$, $\sigma_{ZZ} = 0$). The $J_1 + \langle J_2\rangle$ values are obtained using the optimal control parameter sets of the Hadamard gate from the IDG strategy (blue triangles), QSN strategy (red circles), and TVN strategy (yellow squares). Ten realizations of OU noise $\beta_{\mathrm{OU}}(t)$ with $\sigma_{\mathrm{OU}} = 10^{-3}$ for $\gamma_{\mathrm{OU}}/\omega_0$ equal to (c) 10^{-7}, (d) 10^{-3}, and (e) 10^{-1}.

III. RESULTS AND DEMONSTRATIONS

A. Comparison with the quasistatic-noise method

1. Single-qubit gates

We demonstrate as an example the implementation of single-qubit gates in the presence of TVN using our method. The ideal system Hamiltonian for the qubit is

$$\mathcal{H}_I(t) = \omega_0\frac{Z}{2} + \Omega_X(t)\frac{X}{2}, \tag{9}$$

where X and Z stand for the Pauli matrices, ω_0 is the qubit transition frequency, and $\Omega_X(t)$ is the control field in the X term. The noise Hamiltonian is written as

$$\mathcal{H}_N(t) = \beta_Z(t)\omega_0\frac{Z}{2} + \beta_X(t)\Omega_X(t)\frac{X}{2}. \tag{10}$$

We call $\beta_Z(t)$ the Z noise and $\beta_X(t)$ the X noise and assume that they are independent OU noises with CFs $C_{ZZ}(t_1,t_2) = \sigma_{ZZ}^2\exp\left(-\gamma_{ZZ}|t_1-t_2|\right)$ and $C_{XX}(t_1,t_2) = \sigma_{XX}^2\exp\left(-\gamma_{XX}|t_1-t_2|\right)$ as the form of Eq. (7). We choose the control pulse as a composite sine pulse expressed as

$$\Omega_X(t) = \sum_{k=1}^{k_{\max}} a_k \sin\left(m_k\pi\frac{t}{t_f}\right), \tag{11}$$

where the set of the strengths of the single sine pulses is the control parameter set $[a_k] = [a_1, a_2, \ldots, a_{k_{\max}}]$ and $\{m_k\}$ is a set of integers, chosen depending on the nature of the system Hamiltonians and the target gates as well as the properties of the noise models. For each control pulse, we choose the number of control parameters $k_{\max}$ to range from 8 to 20 in our calculations.

We define below three optimization strategies, namely, the ideal-gate (IDG) strategy, QSN strategy, and TVN strategy. The IDG strategy is to perform the first-step optimization (J_1 optimization) only and to show the performance of an IDG pulse in the presence of noise. The TVN strategy is our proposed method described earlier above, in which the actual γ_{ZZ} and γ_{XX} values are used in the noise CFs of the cost function $\langle J_2\rangle$ for the second-step optimization. The QSN strategy uses the same optimization procedure as the TVN strategy, but with $\gamma_{ZZ} = \gamma_{XX} = 0$ for the noise CFs in the cost function $\langle J_2\rangle$. Thus it is regarded to represent the QSN methods. We choose the gate operation time $t_f = 20/\omega_0$. After the optimizations of the Hadamard gate, we plot the corresponding $J_1 + \langle J_2\rangle$ values obtained from these three strategies versus γ_{ZZ} in Fig. 1(a) for Z noise and versus γ_{XX} in Fig. 1(b) for X noise. For low-frequency (quasistatic) noise ($\gamma_{ZZ} = \gamma_{XX} = 10^{-7}\omega_0$), the performances of the TVN strategy and the QSN strategy are about the same but they are several orders of magnitude better in infidelity $J_1 + \langle J_2\rangle$ value than the IDG strategy, which does not take the noise into account at all. As the noise goes from low frequency to high frequency ($\gamma_{ZZ} = \gamma_{XX} = 10^{-1}\omega_0$), the TVN strategy taking

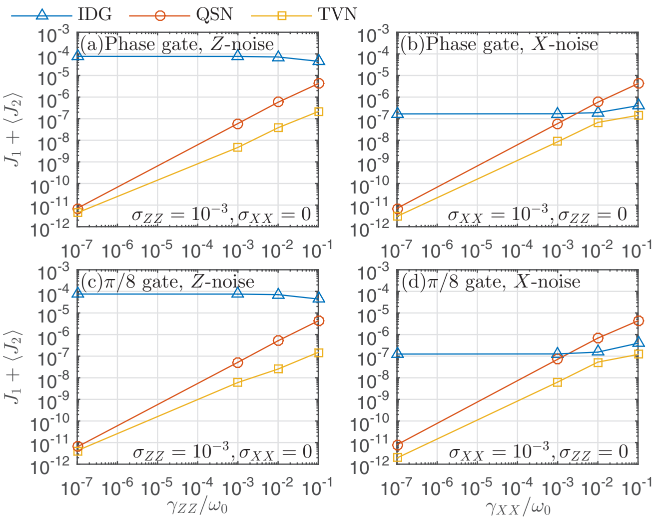

FIG. 2. $J_1 + \langle J_2\rangle$ values versus γ_{ZZ} for Z noise ($\sigma_{ZZ} = 10^{-3}$, $\sigma_{XX} = 0$) and versus γ_{XX} for X noise ($\sigma_{XX} = 10^{-3}$, $\sigma_{ZZ} = 0$) obtained from the IDG strategy (blue triangles), QSN strategy (red circles), and TVN strategy (yellow squares) for the phase gate shown in (a) and (b) respectively, and for the $\pi/8$ gate in (c) and (d), respectively.

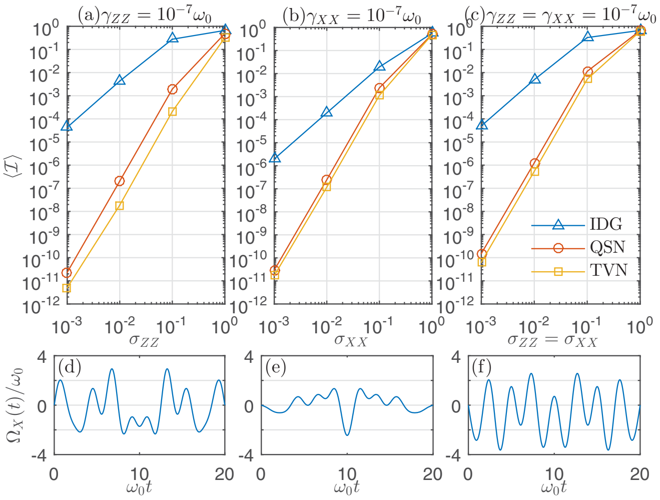

FIG. 3. Robust performance of the Hadamard gate of the IDG strategy (blue triangles), QSN strategy (red circles), and TVN strategy (yellow squares) for low-frequency ($\gamma_{ZZ} = \gamma_{XX} = 10^{-7}\omega_0$) (a) Z noise, (b) X noise, and (c) Z & X noise. The corresponding optimal control pulses of the TVN strategy for (d) Z noise, (e) X noise, and (f) Z & X noise. The number of control parameters $k_{\max} = 10$ for $\Omega_X(t)$ in (d)–(f).

account of the TVN information in the cost function gets better and better (from a factor-level to an order-of-magnitude-level) improvement in $J_1 + \langle J_2\rangle$ values than the QSN strategy, in which noise is assumed to be quasistatic. In addition to the Hadamard gate, we perform calculations for other quantum gates, namely, the phase gate, $\pi/8$ gate and controlled-NOT (CNOT) gate, in the fault-tolerant universal set, in terms of which any unitary operation can be expressed to arbitrary accuracy. The $J_1 + \langle J_2\rangle$ values versus γ_{ZZ} and versus γ_{XX} obtained from the three strategies are shown in Figs. 2(a) and 2(b), respectively, for the phase gate and in Figs. 2(c) and 2(d), respectively, for the $\pi/8$ gate. Their performances are similar to those in Figs. 1(a) and 1(b) of the Hadamard gate. The optimization results for the two-qubit CNOT gate are presented in Sec. III A 2.

Next, we take the optimal control parameter sets of the Hadamard gate from these three strategies to show their robust performance against Z noise, X noise, and Z & X noise at a low frequency ($\gamma_{ZZ} = \gamma_{XX} = 10^{-7}\omega_0$) in Figs. 3(a), 3(b), and 3(c), respectively, and at a high frequency ($\gamma_{ZZ} = \gamma_{XX} = 10^{-1}\omega_0$) in Figs. 4(a), 4(b), and 4(c). For low-frequency noise and for low noise strength ($\sigma_{XX} < 10^{-1}$, $\sigma_{ZZ} < 10^{-1}$), one can see in Fig. 3 that the full-order ensemble average infidelity $\langle\mathcal{I}\rangle$ scales for the IDG strategy as the second power of the noise standard deviation (σ_{ZZ}, σ_{XX}) but scales for the TVN and QSN strategies as the fourth power. This implies that $\langle\mathcal{I}\rangle \cong \langle J_2\rangle$ for the IDG strategy, but the TVN and QSN strategies can nullify the contribution from $\langle J_2\rangle$ for the low-frequency (quasistatic) noise and the dominant contribution to $\langle\mathcal{I}\rangle$ comes from the next-higher-order term, i.e., $\langle\mathcal{I}\rangle \cong \langle O(\tilde{\mathcal{H}}_N^4)\rangle$. In this case, our method, the TVN strategy, still performs slightly better than the QSN strategy. For a gate error (infidelity) less than the error threshold of 10^{-2} of surface codes [49] required for fault-tolerant quantum computation (FTQC), the Hadamard gate of the TVN strategy can be robust to $\sigma_{ZZ} \sim 30\%$ for low-frequency Z noise (i.e., against noise fluctuation, with a standard deviation up to about 30% of $\omega_0/2$), robust to $\sigma_{XX} \sim 20\%$ for X noise [i.e., against noise fluctuation, with a standard deviation up to about 20% of $\Omega_X(t)/2$], and robust to $\sigma_{ZZ} = \sigma_{XX} \sim 10\%$ for Z and X noise as shown in Figs. 3(a), 3(b), and 3(c), respectively. The corresponding optimal control pulses of the TVN strategy are shown in Figs. 3(d), 3(e), and 3(f), respectively.

For the high-frequency noise shown in Fig. 4, the full-order ensemble average infidelity $\langle\mathcal{I}\rangle$ scales as the second power of the noise standard deviation (σ_{ZZ}, σ_{XX}) for all three strategies and noises. This indicates that for high-frequency noise $\langle J_2\rangle$ is not nullified completely but is only minimized. Even in this case, the TVN strategy still has over two orders of magnitude improvement in $\langle\mathcal{I}\rangle$ compared with the IDG strategy and over one order of magnitude improvement compared with the QSN strategy for Z noise at low noise strengths as shown in Fig. 4(a). For $\langle\mathcal{I}\rangle \lesssim 10^{-2}$ less than the FTQC error threshold of the surface codes, the Hadamard gate implemented by our optimal control pulse shown in Fig. 4(d) can be robust to $\sigma_{ZZ} \sim 20\%$ for Z noise. On the other hand, for the high-frequency X noise, $\langle\mathcal{I}\rangle$ obtained by the QSN strategy has values even slightly higher than those by the IDG strategy. The improvement in $\langle\mathcal{I}\rangle$ by the TVN strategy over the other two strategies is less than one order of magnitude. To improve the gate performance, we increase the degrees of freedom for optimization by adding a control term $\Omega_Y(t)Y/2$ and its accompanying Y-noise term $\beta_Y(t)\Omega_Y(t)Y/2$ to the Hamiltonian. We choose, for simplicity, $\gamma_{YY} = \gamma_{XX}$ and $\sigma_{YY} = \sigma_{XX}$ and use the same optimal procedure as for the TVN strategy. The improvement in $\langle\mathcal{I}\rangle$ of the TVN strategy with an additional Y control compared with the TVN strategy is over a half-order of magnitude. As a result, the Hadamard gate with the optimal control pulses of the TVN strategy with an additional Y control shown in Fig. 4(e) can be robust to

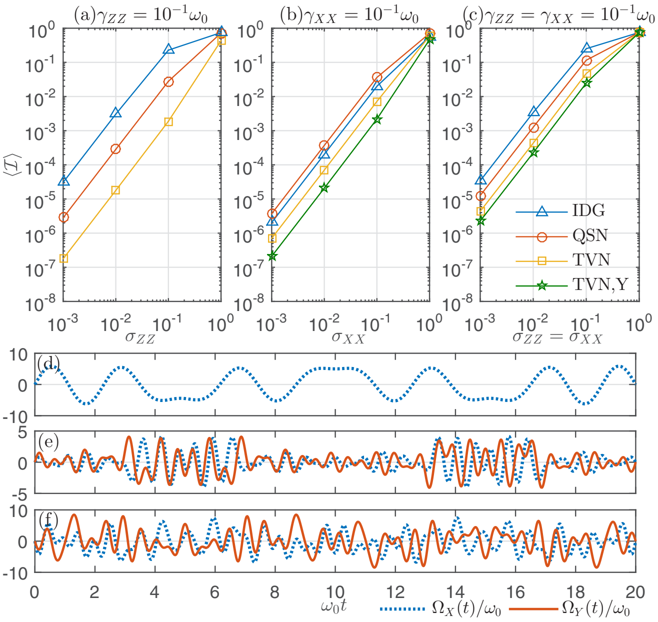

FIG. 4. Robust performance of the Hadamard gate of the IDG strategy (blue triangles), QSN strategy (red circles), and TVN strategy (yellow squares) for high-frequency ($\gamma_{ZZ} = \gamma_{XX} = 10^{-1}\omega_0$) (a) Z noise, (b) X noise, and (c) Z & X noise. For the TVN strategy with an additional Y control (green pentagrams) in (b), $\gamma_{YY} = \gamma_{XX} = 10^{-1}\omega_0$ and $\sigma_{YY} = \sigma_{XX}$, and in (c), $\gamma_{YY} = \gamma_{ZZ} = \gamma_{XX} = 10^{-1}\omega_0$ and $\sigma_{YY} = \sigma_{XX} = \sigma_{ZZ}$. Optimal control pulses of the TVN strategy (d) for Z noise and of the TVN strategy with an additional Y control and accompanying Y noise (e) for X noise and (f) for Z & X noise. The number of control parameters $k_{\max} = 10$ for $\Omega_X(t)$ in (d) and $k_{\max} = 20$ for both $\Omega_X(t)$ and $\Omega_Y(t)$ in (e) and (f).

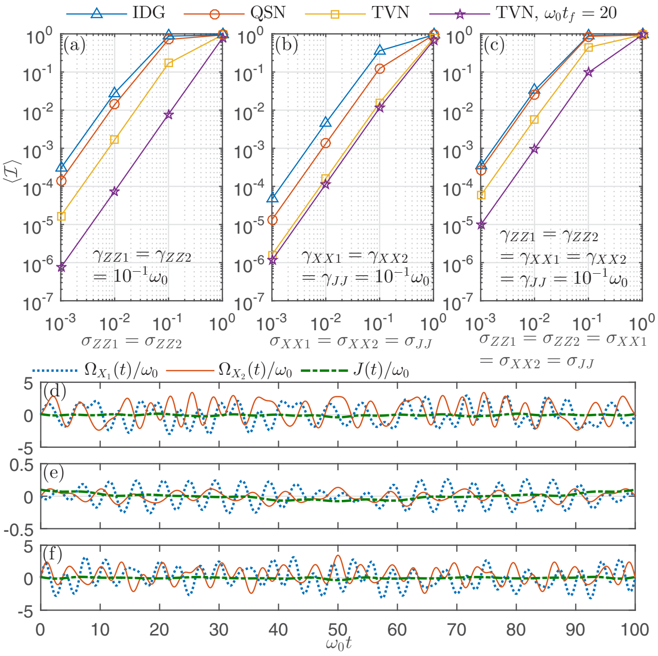

FIG. 5. Robust performance of CNOT gates in the IDG strategy ($\omega_0 t_f = 100$, blue triangles), QSN strategy ($\omega_0 t_f = 100$, red circles), and TVN strategy ($\omega_0 t_f = 100$, yellow squares; and $\omega_0 t_f = 20$, purple pentagrams) for high-frequency ($\gamma_{ZZ1} = \gamma_{ZZ2} = \gamma_{XX1} = \gamma_{XX2} = \gamma_{JJ} = 10^{-1}\omega_0$) (a) Z noise, (b) X & J noise, and (c) Z & X & J noise. The optimal control pulses $\Omega_{X_1}(t)$ (thick dotted blue line), $\Omega_{X_2}(t)$ (thin solid red line), and $J(t)$ (thick dash-dotted green line) of the TVN strategy ($\omega_0 t_f = 100$) for (d) Z noise, (e) X & J noise, and (f) Z & X & J noise. The numbers of control parameters $k_{\max} = 16$, 16, and 8 for $\Omega_{X_1}(t)$, $\Omega_{X_2}(t)$, and $J(t)$, respectively, in (d) and (f); $k_{\max} = 12$, 12, and 6 for $\Omega_{X_1}(t)$, $\Omega_{X_2}(t)$, and $J(t)$, respectively, in (e).

$\sigma_{XX} = \sigma_{YY} \sim 20\%$ for $\langle\mathcal{I}\rangle \lesssim 10^{-2}$. Note that the optimization algorithm seems to find control pulses with stronger strengths to suppress the Z noise but searches weaker control pulses to minimize the X-noise cost function since the system coupling operator term of the X noise is proportional to the control field $\Omega_X(t)$ in our noise model. So for the case with the Z noise and X noise simultaneously present, there is a trade-off in the control pulse strength for the cost function optimization between the Z noise and the X noise. Consequently, the ensemble infidelity of the Z & X noise does not reach a low value as in the case with only Z noise or X noise. Thus one can see in Fig. 4(c) that the improvement in $\langle\mathcal{I}\rangle$ of the TVN strategy over the IDG strategy is just near one order of magnitude, and only a half-order compared with the QSN strategy. A similar trade-off also takes place for the TVN strategy with additional Y control, although it performs slightly better than the TVN strategy with only the $\Omega_X(t)$ control field. Nevertheless, the Hadamard gate implemented with the optimal pulse obtained by the TVN strategy with additional Y control shown in Fig. 4(f) can still be robust to $\sigma_{ZZ} = \sigma_{XX} = \sigma_{YY} \sim 6\%$ for $\langle\mathcal{I}\rangle \lesssim 10^{-2}$.

2. Two-qubit gates

Next, we demonstrate that our method can find control pulses for high-fidelity two-qubit CNOT gate operations in the presence of multiple sources of high-frequency noise. The two-qubit Hamiltonian is chosen as

$$\mathcal{H}_I(t) = \omega_0\frac{Z_1}{2} + \Omega_{X_1}(t)\frac{X_1}{2} + \omega_0\frac{Z_2}{2} + \Omega_{X_2}(t)\frac{X_2}{2} + J(t)\frac{Z_1Z_2}{2}, \tag{12}$$

where Z_j and X_j denote Pauli's matrix operators for qubit j, $\Omega_{X_j}(t)$ is the control field applied to qubit j, and $J(t)$ is the two-qubit coupling strength. We assume that OU noise can be present in each of the five terms, and σ_{ZZ1}, σ_{ZZ2}, σ_{XX1}, σ_{XX2}, and σ_{JJ} are, respectively, the corresponding standard deviation σ_{OU}, and γ_{ZZ1}, γ_{ZZ2}, γ_{XX1}, γ_{XX2}, and γ_{JJ} are, respectively, the corresponding γ_{OU}. We choose the control fields $\Omega_{X_1}(t)$ and $\Omega_{X_2}(t)$ as composite sine pulses and the two-qubit control $J(t)$ as a composite sine pulse with a constant shift.

The robust performance of the CNOT gate using the three strategies for high-frequency ($\gamma_{ZZ1} = \gamma_{ZZ2} = \gamma_{XX1} = \gamma_{XX2} = \gamma_{JJ} = 10^{-1}\omega_0$) Z noise, X & J noise, and Z & X & J noise are shown in Figs. 5(a), 5(b), and 5(c), respectively. The corresponding optimal control pulses of the TVN strategy for an operation time $t_f = 100/\omega_0$ are shown in Figs. 5(d), 5(e), and 5(f), respectively. For $\omega_0 t_f = 100$, our method (the TVN strategy) in the case of Z noise and the case of X & J noise shows a one order of magnitude improvement in $\langle\mathcal{I}\rangle$ values compared with the QSN strategy for low noise strength, but only a half-order improvement in the case of Z & X & J noise. This is because in the case of the Z & X & J noise, there is a trade-off in the control pulse strength for the cost function optimization between the Z noise and the X & J noise, similar to that in the single-qubit case. The robust performance can be improved by reducing the gate operation time t_f, for example, from $t_f = 100/\omega_0$ to $t_f = 20/\omega_0$, to decrease the duration of the influence of the noises. This is shown by the purple pentagrams in Figs. 5(a) and 5(c). In the case of X & J noise in Fig. 5(b), only a slight improvement is observed in the $t_f = 20/\omega_0$ case, because when the operation time decreases, it is hard to make the strengths of the control fields $\Omega_X^j(t)$ and $J(t)$ all low as in the $t_f = 100/\omega_0$ case. For high-frequency noise and for the FTQC error threshold $\langle\mathcal{I}\rangle \lesssim 10^{-2}$ of the surface codes, the CNOT gate with operation time $t_f = 20/\omega_0$ can be robust to $\sigma_{ZZ1} = \sigma_{ZZ2} \sim 10\%$ for Z noise, robust to $\sigma_{XX1} = \sigma_{XX2} = \sigma_{JJ} \sim 10\%$ for X & J noise, and robust to $\sigma_{ZZ1} = \sigma_{ZZ2} = \sigma_{XX1} = \sigma_{XX2} = \sigma_{JJ} \sim 3\%$ for Z & X & J noise by our method.

We describe briefly the computational resources and computation time in our calculations. In the case of Z & X & J noise, we use 40 control parameters in a parameter set to run the two-step optimization for the two-qubit CNOT gate and choose 100 initial random guesses of the parameter sets for the first-step optimization and 10 parameter sets obtained in the first-step optimization as initial guesses for the second-step optimization. We use a total of 60 2-GHz-CPU cores and it takes about 2 days to obtain the control pulses and robust performance calculations in Fig. 5(c). These resources and time spent constructing robust high-fidelity CNOT gates against five sources of high-frequency noise are quite acceptable.

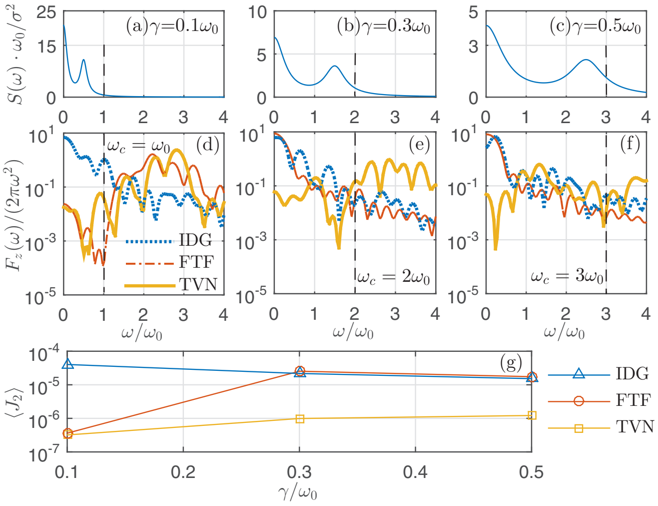

FIG. 6. The behavior of $[F_z(\omega)/(2\pi\omega^2)]$ obtained using the optimal control parameter sets from the IDG strategy (thick dotted blue line), FTF (thin dash-dotted red line) strategy, and TVN strategy (thick solid yellow line) for the noise PSD $S(\omega)$ with (a) $\gamma = 0.1\omega_0$, (b) $\gamma = 0.3\omega_0$, and (c) $\gamma = 0.5\omega_0$ is shown in (d), (e), and (f) respectively. (g) The corresponding $\langle J_2\rangle$ values.

B. Comparison with the filter-transfer-function method

In this subsection, we compare our method with the FTF method [38–40]. The cost function $\langle J_2\rangle$ in Eq. (6) can be transformed to the frequency domain as

$$\langle J_2\rangle = \sum_j \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{d\omega}{\omega^2}S_j(\omega)F_j(\omega), \tag{13}$$

where $S_j(\omega)$ is the noise PSD for the jth noise, and $F_j(\omega)$ is the corresponding filter-transfer function. The cost function of the jth noise for optimization in the FTF method is defined as $A_j \equiv \int_{\omega_L}^{\omega_c} F_j(\omega)d\omega$ [38–40]. The region $[\omega_L,\omega_c]$ of the integration of the cost function A_j is determined by the nonnegligible region of the noise PSD. In order to make comparisons with our method, we use the same form of control pulse, the same number of control parameters, and the same optimal procedure except, for the FTF method, changing the cost function from $J_1 + \langle J_2\rangle$ to $J_1 + A_j$ in the second step of the two-step optimization. We call this procedure the FTF strategy. Then we apply the IDG strategy, FTF strategy, and TVN strategy to find the high-fidelity Hadamard gate for a one-qubit system with single Z noise. To demonstrate the advantage of our method over the FTF method, we choose the noise PSD for Z noise to contain a high-frequency distribution as

$$S(\omega) = \sigma^2\left[\frac{2\gamma}{\gamma^2+\omega^2} + \frac{\gamma}{\gamma^2+(5\gamma-\omega)^2} + \frac{\gamma}{\gamma^2+(5\gamma+\omega)^2}\right], \tag{14}$$

which has two peaks, at $\omega = 0$ and $\omega = 5\gamma$. As the value of γ increases, the dominant distribution associated with the second peak of the PSD $S(\omega)$ moves to a high-frequency region in which the FTF method may not work very effectively. We demonstrate that our method, including the detailed noise PSD distribution in the cost function, can still in this case suppress the gate error coming from $S(\omega)$, a PSD different from that in the OU noise model used previously. The lower limit ω_L of the integral of the cost function A_z for the FTF strategy is chosen to be 0, and the upper limit ω_c is chosen to be $1\omega_0$, $2\omega_0$, and $3\omega_0$ to enclose the dominant distribution of $S(\omega)$ [see Figs. 6(a), 6(b), and 6(c)] for $\gamma = 0.1\omega_0$, $0.3\omega_0$, and $0.5\omega_0$, respectively. For the single Z noise considered here, the infidelity from Eq. (13) is $\langle J_2\rangle = \int_{-\infty}^{\infty} d\omega S(\omega)[F_z(\omega)/(2\pi\omega^2)]$. The improvement of $\langle J_2\rangle$ can be analyzed through the overlap of $S(\omega)$ with $[F_z(\omega)/(2\pi\omega^2)]$ [35]. If the control pulses can make $[F_z(\omega)/(2\pi\omega^2)]$ small in the dominant distribution region of $S(\omega)$, then $\langle J_2\rangle$ can be significantly improved (reduced). We plot $[F_z(\omega)/(2\pi\omega^2)]$ evaluated by the optimal control parameter sets obtained from the above three strategies for three values, $\gamma = 0.1\omega_0$, $0.3\omega_0$, and $0.5\omega_0$ of $S(\omega)$, in Figs. 6(d), 6(e), and 6(f), respectively. The corresponding $\langle J_2\rangle$ values are shown in Fig. 6(g). By taking the case of $\gamma = 0.3\omega_0$ as an example, the function $[F_z(\omega)/(2\pi\omega^2)]$ of the TVN strategy shows apparent drops near the two peaks of the noise PSD at $\omega = 0$ and $\omega = 1.5\omega_0$, but the function for the FTF strategy and the IDG strategy does not. Thus, an improvement of about one order of magnitude in $\langle J_2\rangle$ of the TVN strategy over the other two strategies is observed. In short, as the range of dominant distribution of the PSD increases [e.g., from Fig. 6(a) to Fig. 6(c)], the TVN strategy, including the detailed noise information (CF) in the optimization cost function [35], can suppress the dominant infidelity contribution more effectively than the FTF strategy. Furthermore, the concatenation method is used to construct control pulses against two different noncommuting noises in the FTF method [39,40]. But using the concatenation method to deal with the case of multicontrols, multiple sources of noise, and multiqubits may be very complicated. On the other hand, our method can find robust control pulses for high-fidelity CNOT gates that involve three control knobs and up to five sources of high-frequency noise as demonstrated in Fig. 5.

IV. CONCLUSION

To conclude, our two-step optimization method can provide robust control pulses of high-fidelity quantum gates for stochastic TVN. Besides, our method is quite general and can be applied to different system models, noise models, and noise CFs (PSDs). If the system is very clean, i.e., quantum noise is very weak, the method presented in this paper can be used directly to find control pulses for high-fidelity gates in the presence of time-varying classical noises. For example, the decoherence (dephasing) time is rather long in quantum-dot electron spin qubits in purified silicon [50,51] compared to those in GaAs [52,53], that is, quantum noise coming from the coupling to the environmental spins is very weak, and the dominant gate error is due to the electric classical control noise [51]. We present the results for quantum-dot electron spin qubits in purified silicon elsewhere. Our method will take essential steps toward constructing high-fidelity and robust quantum gates for FTQC in realistic quantum computing systems.

ACKNOWLEDGMENTS

We acknowledge support from the Ministry of Science and Technology of Taiwan under Grant No. 103-2112-M-002-003-MY3, from the National Taiwan University under Grants No. 105R891402 and No. 105R104021, and from the thematic group program of the National Center for Theoretical Sciences, Taiwan.

APPENDIX A: DERIVATION OF EQS. (2)–(4)

We present the derivation of Eqs. (2)–(4) and discuss the role of the extra term ϵ in Eq. (2). Substituting the total system propagator in the Dyson expansion $U(t_f) = U_I(t_f)(I + \Psi_1 + \Psi_2 + \cdots)$ into the infidelity definition $\mathcal{I}$ of Eq. (1), we obtain

$$\mathcal{I} = J_1 - \frac{2}{4^n}\mathrm{Re}\{\mathrm{Tr}[U_T^\dagger U_I(t_f)]^\star \cdot \mathrm{Tr}[U_T^\dagger U_I(t_f)\cdot(\Psi_1+\Psi_2+\cdots)]\} - \frac{1}{4^n}|\mathrm{Tr}[U_T^\dagger U_I(t_f)\cdot(\Psi_1+\Psi_2+\cdots)]|^2. \tag{A1}$$

The first term J_1 on the right-hand side of Eq. (A1) is the gate infidelity for the ideal system defined in Eq. (3). Then we define the error shift matrix U_ϵ of the ideal propagator $U_I(t_f)$ at time t_f from the target gate U_T up to a global phase ϕ as

$$U_I(t_f) = e^{i\phi}U_T(I + U_\epsilon). \tag{A2}$$

Note that when the gate infidelity J_1 for the ideal system is made small, the matrix elements of U_ϵ also become small. Substituting the expression of $U_I(t_f)$ of Eq. (A2) back into Eq. (A1), we obtain

$$\mathcal{I} = J_1 + \left\{-\frac{1}{2^{n-1}}\mathrm{Re}[\mathrm{Tr}(\Psi_1)]\right\} + J_2 + \epsilon(U_\epsilon,\Psi_j) + O\left(\tilde{\mathcal{H}}_N^m, m \geqslant 3\right), \tag{A3}$$

where J_2 is defined in Eq. (4),

$$\begin{aligned}\epsilon(U_\epsilon,\Psi_j) = &-\frac{1}{2^{n-1}}\mathrm{Re}\{\mathrm{Tr}[U_\epsilon(\Psi_1+\Psi_2+\cdots)]\} - \frac{2}{4^n}\mathrm{Re}\{\mathrm{Tr}[U_\epsilon]^\star\cdot\mathrm{Tr}[\Psi_1+\Psi_2+\cdots]\}\\ &-\frac{2}{4^n}\mathrm{Re}\{\mathrm{Tr}[U_\epsilon(\Psi_1+\Psi_2+\cdots)]\cdot\mathrm{Tr}[\Psi_1+\Psi_2+\cdots]^\star\} - \frac{2}{4^n}\mathrm{Re}\{\mathrm{Tr}[U_\epsilon]^\star\cdot\mathrm{Tr}[U_\epsilon(\Psi_1+\Psi_2+\cdots)]\}\\ &-\frac{1}{4^n}|\mathrm{Tr}[U_\epsilon(\Psi_1+\Psi_2+\cdots)]|^2,\end{aligned} \tag{A4}$$

and $O(\tilde{\mathcal{H}}_N^m, m \geqslant 3)$ denotes other higher-order terms without containing U_ϵ. The first-order noise term, $-\mathrm{Re}[\mathrm{Tr}(\Psi_1)]/2^{n-1}$, in Eq. (A3) actually vanishes. Because the noise Hamiltonian $\mathcal{H}_N$ is Hermitian [with $\beta_j(t)$ being real], $\mathrm{Tr}[\mathcal{H}_N(t')]$ is a real number. Thus the first-order term proportional to the real part of $\mathrm{Re}[\mathrm{Tr}(\Psi_1)]$, with $\mathrm{Tr}(\Psi_1) = -i\int_0^{t_f}\mathrm{Tr}[\mathcal{H}_N(t')]dt'$, vanishes. This result, no first-order noise contribution to $\mathcal{I}$, is similar to that in Ref. [32]. This is also the reason why there is no first-order noise contribution to the ensemble average $\langle\mathcal{I}\rangle$ of Eq. (5). Equations (2)–(4) can then be easily obtained from Eq. (A3) with the identification of $\epsilon = \epsilon(U_\epsilon,\Psi_j)$.

We discuss below the properties and the role of $\epsilon = \epsilon(U_\epsilon,\Psi_j)$ in Eq. (2) or in Eq. (A3). The extra contribution $\epsilon = \epsilon(U_\epsilon,\Psi_j)$ to the gate infidelity, with the detailed form shown in Eq. (A4), is related to the error shift matrix U_ϵ and all Dyson expansion terms Ψ_j. As noted earlier, if J_1 is small, then the matrix elements of U_ϵ are also small. Moreover, if the noise strength is not too strong such that $|\Psi_{j+1}| \ll |\Psi_j|$, then the extra contribution $\epsilon = \epsilon(U_\epsilon,\Psi_j)$ is also small. Therefore when running the optimization for a low noise strength, for which the higher-order term $O(\tilde{\mathcal{H}}_N^m, m \geqslant 3)$ becomes negligible (see Appendix B), the extra contribution ϵ can be omitted as J_1 is minimized to a small number. Consequently, one can focus on the optimization of only $J_1 + \langle J_2 \rangle$.

The advantage of introducing J_1 and ϵ in our method is to enable more degrees of freedom in control parameters for optimization. There are actually no J_1 and ϵ contributions to the gate infidelity expression of the robust control method of SUPCODE [20,21] and the filter-transfer-function method [39,40]. In these methods, J_1, or, equivalently, the error shift matrix U_ϵ, is set exactly to 0 by imposing some constraints on the control parameters. In contrast, our method can tolerate some error in U_ϵ and thus have more degrees of freedom in control parameters as long as J_1 and the extra contribution $\langle \epsilon \rangle$ to the gate infidelity $\langle \mathcal{I} \rangle$ are made just smaller than $\langle J_2 \rangle$. This advantage of having more degrees of freedom for optimization plays an important role in finding better control pulses as the number of qubits, the number of controls, and the number of noise sources increase.

APPENDIX B: ESTIMATION OF HIGHER-ORDER CONTRIBUTIONS

Here we estimate the contributions of higher-order terms $O(\tilde{\mathcal{H}}_N^m, m \geqslant 3)$ and discuss when they can be neglected. We express the higher-order terms as $O(\tilde{\mathcal{H}}_N^m, m \geqslant 3) = \sum_{p\geqslant 3} J_p$, where J_p denotes the pth-order noise term of the gate infidelity. Detailed forms of the first two lowest-order terms in $O(\tilde{\mathcal{H}}_N^m, m \geqslant 3)$ are

$$J_3 = -\frac{1}{2^{n-1}}\mathrm{Re}[\mathrm{Tr}(\Psi_3)] - \frac{2}{4^n}\mathrm{Re}\{\mathrm{Tr}(\Psi_1)\mathrm{Tr}(\Psi_2)^\star\}, \tag{B1}$$

$$J_4 = -\frac{1}{2^{n-1}}\mathrm{Re}[\mathrm{Tr}(\Psi_4)] - \frac{1}{4^n}|\mathrm{Tr}(\Psi_2)|^2 - \frac{2}{4^n}\mathrm{Re}\{\mathrm{Tr}(\Psi_1)\mathrm{Tr}(\Psi_3)^\star\}, \tag{B2}$$

where

$$\Psi_q = (-i)^q \int_0^{t_f} dt_1 \int_0^{t_1} dt_2 \dots \times \int_0^{t_{q-1}} dt_q \tilde{\mathcal{H}}_N(t_1)\tilde{\mathcal{H}}_N(t_2)\dots\tilde{\mathcal{H}}_N(t_q) \tag{B3}$$

is the qth-order Dyson expansion term. To make an estimation of the magnitude of Ψ_q, we take the Z-noise model for the single-qubit gate operations in Sec. III A 1 as an example. Substituting the noise Hamiltonian $\tilde{\mathcal{H}}_N(t) = \beta_Z(t)R_Z(t)$ with $R_Z(t) = U_I^\dagger(t)[\omega_0 Z/2]U_I(t)$ in the interaction picture into Ψ_q, we obtain

$$\Psi_q = (-i)^q \int_0^{t_f} \omega_0 dt_1 \int_0^{t_1} \omega_0 dt_2 \dots \int_0^{t_{q-1}} \omega_0 dt_q \times \{\beta_Z(t_1)\beta_Z(t_2)\dots\beta_Z(t_q)\}\{\bar{R}_Z(t_1)\bar{R}_Z(t_2)\dots\bar{R}_Z(t_q)\}, \tag{B4}$$

where $\bar{R}_Z(t) = U_I^\dagger(t)[Z/2]U_I(t)$. Since $U_I(t)$ is unitary, its matrix elements $|U_{I,jk}(t)| \leqslant 1$ for all j and k. Consequently, $|\bar{R}_{Z,jk}(t)| < 1$ for all j and k, so $|\{\bar{R}_Z(t_1)\bar{R}_Z(t_2)\dots\bar{R}_Z(t_q)\}_{jk}| < 1$ for all j and k. Taking the strength of $\beta_Z(t)$ to be about its standard deviation σ_{ZZ}, we estimate the noise strength contribution to be $|\{\beta_Z(t_1)\beta_Z(t_2)\dots\beta_Z(t_q)\}| \approx (\sigma_{ZZ})^q$. The time integral contribution $\{\int_0^{t_f}\omega_0 dt_1 \int_0^{t_1}\omega_0 dt_2 \dots \int_0^{t_{q-1}}\omega_0 dt_q\}$ can be estimated to be $\sim(\omega_0 t_f)^q/q!$. By combining the above estimations, the magnitude of $|\Psi_{q,jk}|$ is of the order of $\sim(\omega_0 t_f \sigma_{ZZ})^q/q!$. Then substituting the estimated value of $|\Psi_{q,jk}|$ into J_2 in Eq. (4), J_3 in Eq. (B1), and J_4 in Eq. (B2), we obtain the magnitude ratios $J_3/J_2 \sim (\omega_0 t_f \sigma_{ZZ})/3$ and $J_4/J_2 \sim (\omega_0 t_f \sigma_{ZZ})^2/12$. The single-qubit gate operation time in Sec. III A 1 is $\omega_0 t_f = 20$. If we choose the noise fluctuation $\sigma_{ZZ} = 10^{-3}$, then the ratio $J_3/J_2 \sim (6 \times 10^{-3})$ and $J_4/J_2 \sim (3 \times 10^{-5})$, and thus the higher-order terms $O(\tilde{\mathcal{H}}_N^m, m \geqslant 3)$ can be safely neglected. If, however, $\sigma_{ZZ} \sim 10^{-1}$, then $\omega_0 t_f \sigma_{ZZ} \sim 2$. In this case, $J_3/J_2 \sim 2/3$ and $J_4/J_2 \sim 1/3$, so the higher-order terms $O(\tilde{\mathcal{H}}_N^m, m \geqslant 3)$ cannot be neglected. Comparing our estimation with the results of the full-Hamiltonian simulation, one finds that the ensemble average of the gate infidelity $\langle \mathcal{I} \rangle$ of the IDG strategy scales as the second power of σ_{ZZ} (because $\langle J_2 \rangle$ dominates) for small σ_{ZZ} until $\sigma_{ZZ} \sim 10^{-1}$ in Fig. 3(a) for low-frequency noise $\gamma_{ZZ} = 10^{-7}\omega_0$ and in Fig. 4(a) for high-frequency noise $\gamma_{ZZ} = 10^{-1}\omega_0$. This is consistent with our estimation. In other words, if σ_{ZZ} is considerably smaller than 10^{-1}, $O(\tilde{\mathcal{H}}_N^m, m \geqslant 3)$ can be ignored. Therefore, even in the case where the full-Hamiltonian simulation is not available, we can use this estimation method to determine the criterion for neglecting the higher-order terms $O(\tilde{\mathcal{H}}_N^m, m \geqslant 3)$.

---

[1] K. Khodjasteh and L. Viola, Phys. Rev. Lett. **102**, 080501 (2009); Phys. Rev. A **80**, 032314 (2009); K. Khodjasteh, D. A. Lidar, and L. Viola, Phys. Rev. Lett. **104**, 090501 (2010); K. Khodjasteh, H. Bluhm, and L. Viola, Phys. Rev. A **86**, 042329 (2012).

[2] J. R. West, D. A. Lidar, B. H. Fong, and M. F. Gyure, Phys. Rev. Lett. **105**, 230503 (2010).

[3] A. M. Souza, G. A. Álvarez, and D. Suter, Phys. Rev. A **86**, 050301 (2012).

[4] F. F. Fanchini, J. E. M. Hornos, and R. d. J. Napolitano, Phys. Rev. A **76**, 032319 (2007); F. F. Fanchini, R. d. J. Napolitano, B. Çakmak, and A. O. Caldeira, *ibid.* **91**, 042325 (2015).

[5] A. Z. Chaudhry and J. Gong, Phys. Rev. A **85**, 012315 (2012).

[6] X. Xu, Z. Wang, C. Duan, P. Huang, P. Wang, Y. Wang, N. Xu, X. Kong, F. Shi, X. Rong, and J. Du, Phys. Rev. Lett. **109**, 070502 (2012).

[7] J. Clausen, G. Bensky, and G. Kurizki, Phys. Rev. Lett. **104**, 040401 (2010); Phys. Rev. A **85**, 052105 (2012).

[8] M. Wenin and W. Pötz, Phys. Rev. B **78**, 165118 (2008).

[9] P. Rebentrost, I. Serban, T. Schulte-Herbrüggen, and F. K. Wilhelm, Phys. Rev. Lett. **102**, 090401 (2009).

[10] B. Hwang and H.-S. Goan, Phys. Rev. A **85**, 032321 (2012); J.-S. Tai, K.-T. Lin, and H.-S. Goan, *ibid.* **89**, 062310 (2014); S.-Y. Huang and H.-S. Goan, *ibid.* **90**, 012318 (2014); Y. Chou, S.-Y. Huang, and H.-S. Goan, *ibid.* **91**, 052315 (2015).

[11] R. Tycko, Phys. Rev. Lett. **51**, 775 (1983).

[12] M. H. Levitt, Prog. Nucl. Magn. Reson. Spectrosc. **18**, 61 (1986).
[13] S. Wimperis, J. Magn. Reson. Ser. A **109**, 221 (1994).
[14] H. K. Cummins, G. Llewellyn, and J. A. Jones, Phys. Rev. A **67**, 042308 (2003).
[15] K. R. Brown, A. W. Harrow, and I. L. Chuang, Phys. Rev. A **70**, 052318 (2004).
[16] L. M. K. Vandersypen and I. L. Chuang, Rev. Mod. Phys. **76**, 1037 (2005).
[17] T. Ichikawa, M. Bando, Y. Kondo, and M. Nakahara, Phys. Rev. A **84**, 062311 (2011); M. Bando, T. Ichikawa, Y. Kondo, and M. Nakahara, J. Phys. Soc. Jpn. **82**, 014004 (2013).
[18] T. Ichikawa, U. Güngördü, M. Bando, Y. Kondo, and M. Nakahara, Phys. Rev. A **87**, 022323 (2013).
[19] J. T. Merrill and K. R. Brown, in *Quantum Information and Computation for Chemistry*, edited by S. Kais (John Wiley & Sons, New York, 2014), pp. 241–294.
[20] X. Wang, L. S. Bishop, J. P. Kestner, E. Barnes, K. Sun, and S. Das Sarma, Nat. Commun. **3**, 997 (2012).
[21] J. P. Kestner, X. Wang, L. S. Bishop, E. Barnes, and S. Das Sarma, Phys. Rev. Lett. **110**, 140502 (2013).
[22] X. Wang, L. S. Bishop, E. Barnes, J. P. Kestner, and S. Das Sarma, Phys. Rev. A **89**, 022310 (2014).
[23] X. Wang, F. A. Calderon-Vargas, M. S. Rana, J. P. Kestner, E. Barnes, and S. Das Sarma, Phys. Rev. B **90**, 155306 (2014).
[24] X. Rong, J. Geng, Z. Wang, Q. Zhang, C. Ju, F. Shi, C.-K. Duan, and J. Du, Phys. Rev. Lett. **112**, 050503 (2014).
[25] X.-C. Yang and X. Wang, Sci. Rep. **6**, 28996 (2016).
[26] C. Chen, D. Dong, R. Long, I. R. Petersen, and H. A. Rabitz, Phys. Rev. A **89**, 023402 (2014).
[27] D. Dong, C. Chen, B. Qi, I. R. Petersen, and F. Nori, Sci. Rep. **5**, 7873 (2015).
[28] D. Dong, C. Wu, C. Chen, B. Qi, I. R. Petersen, and F. Nori, Sci. Rep. **6**, 36090 (2016).
[29] J.-S. Li and N. Khaneja, Phys. Rev. A **73**, 030302 (2006).
[30] J. H. Lee, E. Montano, I. H. Deutsch, and P. S. Jessen, Nat. Commun. **4**, 2027 (2013).
[31] E. Barnes, X. Wang, and S. Das Sarma, Sci. Rep. **5**, 12685 (2015).
[32] D. Daems, A. Ruschhaupt, D. Sugny, and S. Guérin, Phys. Rev. Lett. **111**, 050404 (2013).
[33] R. L. Kosut, M. D. Grace, and C. Brif, Phys. Rev. A **88**, 052326 (2013).
[34] B. E. Anderson, H. Sosa-Martinez, C. A. Riofrío, I. H. Deutsch, and P. S. Jessen, Phys. Rev. Lett. **114**, 240401 (2015).
[35] G. Gordon, G. Kurizki, and D. A. Lidar, Phys. Rev. Lett. **101**, 010403 (2008).
[36] A. Ruschhaupt, X. Chen, D. Alonso, and J. Muga, New J. Phys. **14**, 093040 (2012).
[37] X.-J. Lu, X. Chen, A. Ruschhaupt, D. Alonso, S. Guérin, and J. G. Muga, Phys. Rev. A **88**, 033406 (2013).
[38] T. J. Green, J. Sastrawan, H. Uys, and M. J. Biercuk, New J. Phys. **15**, 095004 (2013).
[39] A. Soare, H. Ball, D. Hayes, J. Sastrawan, M. C. Jarratt, J. J. McLoughlin, X. Zhen, T. J. Green, and M. J. Biercuk, Nat. Phys. **10**, 825 (2014).
[40] H. Ball and M. J. Biercuk, EPJ Quantum Technol. **2**, 11 (2015).
[41] G. A. Paz-Silva and L. Viola, Phys. Rev. Lett. **113**, 250501 (2014).
[42] W. D. Oliver, Nat. Phys. **10**, 794 (2014).
[43] C. Kabytayev, T. J. Green, K. Khodjasteh, M. J. Biercuk, L. Viola, and K. R. Brown, Phys. Rev. A **90**, 012316 (2014).
[44] F. J. Dyson, Phys. Rev. **75**, 486 (1949).
[45] J. A. Nelder and R. Mead, Comput. J. **7**, 308 (1965).
[46] S. Finch and Y. Yt, *Ornstein-Uhlenbeck Process* (Citeseer X, 2004); http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.299.744
[47] A. V. Kuhlmann, J. Houel, A. Ludwig, L. Greuter, D. Reuter, A. D. Wieck, M. Poggio, and R. J. Warburton, Nat. Phys. **9**, 570 (2013).
[48] Y. Li, N. Sinitsyn, D. L. Smith, D. Reuter, A. D. Wieck, D. R. Yakovlev, M. Bayer, and S. A. Crooker, Phys. Rev. Lett. **108**, 186603 (2012).
[49] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, Phys. Rev. A **86**, 032324 (2012).
[50] M. Veldhorst, J. C. C. Hwang, C. H. Yang, A. W. Leenstra, B. de Ronde, J. P. Dehollain, J. T. Muhonen, F. E. Hudson, K. M. Itoh, A. Morello, and A. S. Dzurak, Nat. Nanotech. **9**, 981 (2014).
[51] M. Veldhorst, C. H. Yang, J. C. C. Hwang, W. Huang, J. P. Dehollain, J. T. Muhonen, S. Simmons, A. Laucht, F. E. Hudson, K. M. Itoh, A. Morello, and A. S. Dzurak, Nature **526**, 410 (2015).
[52] J. R. Petta, A. C. Johnson, J. M. Taylor, E. A. Laird, A. Yacoby, M. D. Lukin, C. M. Marcus, M. P. Hanson, and A. C. Gossard, Science **309**, 2180 (2005).
[53] F. H. L. Koppens, C. Buizert, K. J. Tielrooij, I. T. Vink, K. C. Nowack, T. Meunier, L. P. Kouwenhoven, and L. M. K. Vandersypen, Nature **442**, 766 (2006).