Online Optimal Operation of Parallel Voltage Source Inverters Using Partial Information

Kyriakos G. Vamvoudakis, *Senior Member, IEEE*, João P. Hespanha *Fellow Member, IEEE*

***Abstract*—In this paper, a novel optimal learning algorithm for partially unknown voltage source inverters (VSIs) operating in parallel is presented. The algorithm designs game-theory based distributed controllers to provide the appropriate working voltage magnitude and frequency at the load by converting DC voltage to AC voltage at the parallel VSIs. It takes advantage of information from the neighboring low pass $L-C$ filters to improve harmonic distortion and guarantee equal sharing of the load current across the VSIs while avoiding current circulation during transient and ensuring stability and robustness. It builds upon the ideas of approximate dynamic programming (ADP) and uses only partial information of the system and the exosystem, which is connected only to some of the VSIs. The proposed framework was tested in simulations to show its effectiveness.**

***Index Terms*—Parallel VSIs, game theory, learning.**

I. Introduction

With the integration of renewable energy sources like photovoltaic system, wind turbine, fuel cells, batteries, the grid became more vulnerable to islanding phenomenon, as the renewable sources keep energizing the distribution line, irrespective of its connection to the utility grid. Synchronization will be crucial for all the connected inverters to adapt their behavior in any condition. In addition, the transient responses are forfeited and especially when the load is nonlinear or there are sudden-load applications or removals, the harmonics of the load current cannot be split equally. One solution for generating high-power outputs with low-power inputs is to place a number of electronic VSIs in parallel to enhance reliability. A key issue in parallel operation is the distribution of the load current. In an inverter parallel system, the amplitudes and phases of the output voltages of all inverters should strictly be equal to each other to guarantee that each inverter shares the same load current. Otherwise, current circumfluence and overload of some inverters in the inverter parallel system may arise.

Control systems for power systems [12] and specifically for VSIs usually use PI control [20], deadbeat control [10], a combination of them [4], [17] delta modulation, hysteresis control [23], or complex multivariable feedback control [8], [14], [15], [24]. In practice, it is impossible to know exactly all the parameters of the filter connected to the inverter since most of them change with time. The most commonly used control method for improving power sharing in parallel inverters is the voltage/frequency droop control method. This method is better understood in the phasor domain, where active and reactive power are used to adjust the frequency and amplitude of the output voltage. Specifically the authors in [3] control the frequency and the amplitude by imitating the voltage source with a complex finite-output impedance, while the authors in [7] enforced the output impedance to be resistive. Extending the work of [3] and [7] the authors in [22] considered separate cases for pure inductive, pure resistive and complex impedances, and for the latter, they showed that conventional droop control does not achieve balanced power sharing and, for that reason, they proposed a novel droop controller. In [18], [19] the authors used ideas from coupled oscillators, power systems and multi-agent systems to examine synchronization and power sharing in droop-controlled inverters operating in parallel. A distributed control for pulse-width modulation inverters in parallel operation is proposed in [9] to feed an asynchronous motor. Recently the authors of [13] proposed a distributed droop control topology that utilizes only internal filter variables of the VSI to control its operation but without performance or stability guarantees. The authors in [6] proposed an offline sparsity promoting optimal control approach by designing a centralized controller based on slow coherency. Another interesting approach for controlling parallel VSIs is to use distributed Model Predictive Control (MPC). In [16], the authors propose coordination mechanisms based on Lagrangians. But, generally solving the MPC optimization problem and deriving the manipulated variable is computationally

Manuscript received February 23, 2016; revised July 28, 2016 and September 26, 2016; accepted October 22, 2016. This work was supported by ARO grant W911NF-09-D-0001.
K. G. Vamvoudakis, is with the Department of Aerospace and Ocean Engineering, Virginia Tech, Blacksburg, VA 24061-0203 USA, e-mail: kyriakos@vt.edu.
J. P. Hespanha is with the Center for Control, Dynamical-systems and Computation, University of California, Santa Barbara, CA 93106-9560 USA, e-mail: hespanha@ece.ucsb.edu.

challenging. Other MPC problems arise in cases where the models are not known very well, because of strong model dependence.

Contributions: The contributions of the present paper are as follows. First, in order to achieve minimization of the currents' circulation and reduction of the output voltage errors, we employ game-theoretic coupled minimizations that consider *local* performance indices that depend on the voltages and the currents of the local $L-C$ filters in the VSIs. The proposed framework is based on a separate actor/critic approximator for each inverter that does not require knowledge of the values of the load and parasitic resistances and capacitances of the VSIs which in most cases are hard to know precisely. By incorporating local saturated input voltages in the performance index, we avoid the need of pure pulse-width modulation control methods and phasor domain analysis. The main advantage of an analysis that does not require the phasor domain is that one can control transients much faster than the fundamental period. To achieve equal current sharing amongst the inverters, we use appropriate user-defined matrices in the performance indices. Finally there is only the need of a neighborhood signal line to share the information on the current and voltage. This improves reliability with respect to existing droop-control or other current sharing methods that need information about the highest or average current.

Notation: $\mathbb{R}^+$ denotes the set $\{x \in \mathbb{R} : x > 0\}$, $\|.\|_{\mathcal{P}}$ denotes the Euclidean norm weighted by a matrix $\mathcal{P}$. We write $\underline{\lambda}(M)$ $(\bar{\lambda}(M))$ for the minimum (maximum) eigenvalue of matrix M.

II. Problem Formulation

The goal of the present paper is the design of a game-based computational intelligent algorithm to synchronize a network described by a set $\mathcal{N} := \{1, 2, \ldots, N\}$ of VSIs connected in parallel. Specifically, we want to convert the DC sources of $V_{i\text{dc}}$ Volts $\forall i \in \mathcal{N}$ to an AC Voltage V_{iL} with magnitude equal to $\sqrt{2} \times 110V$ and $60Hz$ frequency i.e. $V_{iL} = \sqrt{2}V\cos(377t), \forall i \in \mathcal{N}$ where V is the rms value of the phase voltage at the bus (taken as 110 Volts) for each VSI. We model the communication, employed for data exchange among the inverters, by a fixed topology graph $\mathcal{G} = (\mathcal{V}_{\mathcal{G}}, \mathcal{E}_{\mathcal{G}})$ consisting of a nonempty finite set of N nodes $\mathcal{V}_{\mathcal{G}} = \{n_1, ..., n_N\}$ that represent the inverters and a set of edges or arcs $\mathcal{E}_{\mathcal{G}} \subseteq \mathcal{V}_{\mathcal{G}} \times \mathcal{V}_{\mathcal{G}}$ that represent inter-inverter information exchange links. We assume the graph is simple, e.g. it has no repeated edges and $(n_i, n_i) \notin \mathcal{E}_{\mathcal{G}}, \forall i \in \mathcal{N}$, i.e. no self-loops. The set of neighbors of a node n_i is denoted by $\mathcal{N}_i = \{n_j : (n_j, n_i) \in \mathcal{E}_{\mathcal{G}}\}$, and consists of the set of nodes with arcs incoming to n_i. The number of neighboring inverters of inverter i is denoted by $d_i := |\mathcal{N}_i|$. A structure for parallel VSIs is shown in Figure 2. We denote by Z_i, $\forall i \in \mathcal{N}$ the impedance of the cable between each inverter and the AC Bus connection point. For simplicity in the analysis, we assume that the line impedances Z_i, $\forall i \in \mathcal{N}$ are equal to zero, but in the simulations, we do consider nonzero line impedances to verify the efficacy of the proposed approach. The diagram shown in Figure 1, presents each inverter filtered with an $L-C$ filter (a second order filter giving -40 dB/decade attenuation) in order to reduce the switching harmonics entering the distribution network.

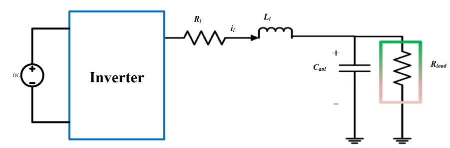


Fig. 1. Schematic representation of each VSI in the network connected to a linear load.

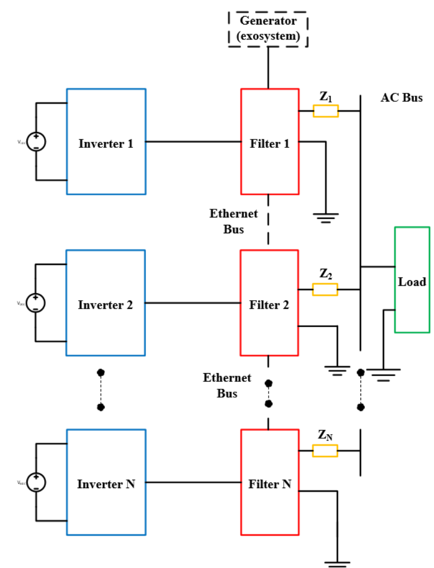


Fig. 2. Schematic representation of a network of VSIs connected in parallel without impedance connection.

To determine a state-space description of the system, we employ Kirchoff's Voltage and Current Laws. In the state space model, the state is $x_i \equiv \begin{bmatrix} i_i & V_{Ci} \end{bmatrix}^T$ and the output V_{iL} is the voltage across the capacitance C_{ap_i} as shown in Figure 1, $\dot{x}_i = A_i x_i + B_i u_i \equiv \begin{bmatrix} -\frac{R_i}{L_i} & -\frac{1}{L_i} \\ \frac{1}{C_{\text{ap}_i}} & -\frac{1}{C_{\text{ap}_i} Z_{\text{load}}} \end{bmatrix} x_i + \begin{bmatrix} \frac{1}{L_i} \\ 0 \end{bmatrix} u_i$, $y_i = C_i^T x_i \equiv \begin{bmatrix} 0 & 1 \end{bmatrix} x_i$.

Remark 1: For simplicity, in the analysis below we consider a pure linear load $Z_{\text{load}} \equiv R_{\text{load}}$. The analysis and the algorithm work for any linear and nonlinear load, as we shall see in the simulation section. This is explained by the fact that our proposed algorithm considers the A_i matrices to be completely unknown. □

To provide the desired frequency and magnitude, we use the following marginally stable exosystem that will be connected to some of the VSIs and will be initialized at the right amplitude: $\dot{Z} \equiv A_Z Z = \begin{bmatrix} 0 & \omega_0 \\ -\omega_0 & 0 \end{bmatrix} Z,\ z(0) = \begin{bmatrix} \sqrt{2}\,110 \\ 0 \end{bmatrix},\ y_0 \equiv C_Z^T Z = \begin{bmatrix} 1 & 0 \end{bmatrix} Z$, where $\omega_0 \in \mathbb{R}^+$ is the desired AC frequency. To design the local criteria to be optimized, we write the output error $e_{\mathrm{rr}i} \in \mathbb{R}$ for all $i \in \mathcal{N}$ that we would like to eventually drive to zero as, $e_{\mathrm{rr}i} = g_i(y_i - y_0) + \sum_{j\in\mathcal{N}_i}(y_i - y_j) = (d_i + g_i)y_i - g_i y_0 - \sum_{j\in\mathcal{N}_i} y_j$, where the so called pinning gains $g_i \geqslant 0$ are nonzero for one or more nodes that are coupled directly to the exosystem. To design distributed controllers, we need to design local performance indices that map the input voltages to the interval $[-V_{\mathrm{idc}}, V_{\mathrm{idc}}]$, and also provide optimal synchronization by exchanging currents and voltages only between the local $L - C$ filters.

We start with the performance index of the form,

$$J_i = \int_0^\infty \big(x_i^T Q_i x_i + R_s(u_i) + \sum_{j\in\mathcal{N}_i} R_s(u_j) + e_{\mathrm{rr}i}^T Q_{ri} e_{\mathrm{rr}i}\big)dt, \forall i \in \mathcal{N} \quad (1)$$

where $Q_i \geq 0, Q_{ri} \geq 0$ are user defined non-negative matrices of appropriate dimensions and $R_s(u_i) = 2\int_0^{u_i} (\theta^{-1}(v))^T dv > 0,\ \forall u_i$ with $v \in \mathbb{R}$ and $\theta(\cdot)$ is a continuous one to one real analytic integrable function of class $C_\mu,\ \mu \geqslant 1$ used to map the interval $[-V_{\mathrm{idc}}, V_{\mathrm{idc}}]$ on to $\mathbb{R}$. This function should be selected to greatly penalize the values of u_i close to the boundaries of $[-V_{\mathrm{idc}}, V_{\mathrm{idc}}]$. A typical choice for that is, $R_s(u_i)$ is $R_s(u_i) = 2\int_0^{u_i}(V_{\mathrm{idc}} \tanh^{-1}(\frac{v}{V_{\mathrm{idc}}}))^T dv$. The term $x_i^T Q_i x_i$ constraints the states (i.e. currents and voltages of the $L - C$ filter) to be small to encourage a smooth response and small current circulation, and the term $e_{\mathrm{rr}i}^T Q_{ri} e_{\mathrm{rr}i}$ promotes exosystem tracking. To avoid a conflict between these two terms, we follow the approach of [1] and restrict the matrix Q_i to be of the form,

$$Q_i = [I - \frac{1}{(d_i + g_i)} C_i (C_i^T C_i)^{-1} C_i^T]^T \hat{Q}_i [I - \frac{1}{(d_i + g_i)} C_i (C_i^T C_i)^{-1} C_i^T] + (d_i + g_i)^2 C_i Q_{ri} C_i^T, \quad (2)$$

for some arbitrary non-negative definite matrix $\hat{Q}_i$. As shown in [1] this choice of Q_i guarantees that $x_i^T Q_i x_i = \hat{x}_i \hat{Q}_i \hat{x}_i$, where $\hat{x}_i = \mathrm{P} x_i$ for a linear projection P with the property that the tracking error is the same for x_i and $\hat{x}_i$. Thus, by penalizing the $\hat{x}_i$ component of x_i, we are not compromising the system's ability to achieve a zero steady state error.

The next step is to write the tracking problem as a regulator problem following [1]. For that reason we shall use an augmented system $\tilde{x}_i := \begin{bmatrix} x_i \\ \mathrm{Col}_{[x_j]_{j\in\mathcal{N}_i}} \\ g_i Z \end{bmatrix}$ with $\mathrm{Col}_{[\cdot_j]_{j\in\mathcal{N}_i}}$ a stacked column vector of the local neighborhood, i.e. $\mathrm{Col}_{[x_j]_{j\in\mathcal{N}_i}} = \{x_j | j \in \mathcal{N}_i\}$.

The augmented dynamics are given by $\forall i \in \mathcal{N}$,

$$\dot{\tilde{x}}_i = \mathrm{diag}\,(A_i, [A_j]_{j\in\mathcal{N}_i}, g_i A_Z)\tilde{x}_i + \begin{bmatrix} \begin{bmatrix} \frac{1}{L_i} \\ 0 \end{bmatrix} u_i \\ \mathrm{Col}_{\left[\begin{bmatrix} \frac{1}{L_j} \\ 0 \end{bmatrix} u_j\right]_{j\in\mathcal{N}_i}} \\ 0_2 \end{bmatrix} \quad (3)$$

with $\mathrm{diag}\big(A_i, [A_j]_{j\in\mathcal{N}_i}, g_i A_Z\big)$ a block diagonal matrix of the local neighborhood. This allows us to re-write (1) with Q_i as in (2) as follows,

$$J_i = \int_0^\infty \big(R_s(u_i) + \sum_{j\in\mathcal{N}_i} R_s(u_j) + \tilde{x}_i^T \tilde{Q}_i \tilde{x}_i\big)dt,\ \forall i \in \mathcal{N}, \quad (4)$$

where $\tilde{Q}_i$ is given at the top of next page (i.e. (5)) with $\mathcal{R}_{[\cdot]_{j\in\mathcal{N}_i}}$ an appended row vector with state information from the local neighborhood. We are interested in finding Nash equilibria policies $u_i^*,\ \forall i \in \mathcal{N}$ for the N-player game in the sense that,

$$J_i(\tilde{x}_i(0); u_i^*, u_{\mathcal{N}_i}^*) \leqslant J_i(\tilde{x}_i(0); u_i, u_{\mathcal{N}_i}^*), \forall u_i, i \in \mathcal{N}, \quad (6)$$

where $J_i(\tilde{x}_i(0); u_i, u_{\mathcal{N}_i}) := \int_0^\infty \big(R_s(u_i) + \sum_{j\in\mathcal{N}_i} R_s(u_j) + \tilde{x}_i^T \tilde{Q}_i \tilde{x}_i\big)d\tau$ subject to the dynamics (3) and given bounded inputs inside the interval $[-V_{\mathrm{idc}}, V_{\mathrm{idc}}]$.

Remark 2: Note that, one can convert easily the performance (4) to consider a minmax optimization ($\mathcal{H}_\infty$ control) in the form of $J_i = \int_0^\infty \big(R_s(u_i) + \sum_{j\in\mathcal{N}_i} R_s(u_j) - \gamma_i^2 \|w_i\|^2 + \tilde{x}_i^T \tilde{Q}_i \tilde{x}_i\big)dt$ with w_i a disturbance input in every inverter and $\gamma_i \geqslant \gamma_i^* \geqslant 0$ with γ_i^* the smallest γ_i that stabilizes the system. The development that follows, can be easily extended in this case following [2]. □

III. Game-Based Distributed Optimization

The equilibria conditions (6) can be expressed by N coupled optimizations,

$$J_i(\tilde{x}_i(0); u_i^*, u_{\mathcal{N}_i}^*) = \min_{u_i} J_i(\tilde{x}_i(0), u_i, u_{\mathcal{N}_i}^*),\ \forall i \in \mathcal{N} \quad (7)$$

which share the dynamics (3) and the Hamiltonians,

$$H_i(\tilde{x}_i, \rho_i, u_i, u_{\mathcal{N}_i}) := \rho_i^T \Bigg(\mathrm{diag}\,(A_i, [A_j]_{j\in\mathcal{N}_i}, g_i A_Z)\tilde{x}_i + \begin{bmatrix} \begin{bmatrix} \frac{1}{L_i} \\ 0 \end{bmatrix} u_i \\ \mathrm{Col}_{\left[\begin{bmatrix} \frac{1}{L_j} \\ 0 \end{bmatrix} u_j\right]_{j\in\mathcal{N}_i}} \\ 0_2 \end{bmatrix} \Bigg) + \big(R_s(u_i) + \sum_{j\in\mathcal{N}_i} R_s(u_j) + \tilde{x}_i^T \tilde{Q}_i \tilde{x}_i\big), \forall i \in \mathcal{N},$$

$$\tilde{Q}_i := \begin{bmatrix} Q_i & -(d_i+g_i)\mathcal{R}[Q_iC_i(C_i^TC_i)^{-1}C_j^T]_{j\in\mathcal{N}_i} & -g_i(d_i+g_i)Q_iC_i(C_i^TC_i)^{-1}C_Z^T \\ -(d_i+g_i)\mathrm{Col}[C_j(C_i^TC_i)^{-1}C_i^TQ_i]_{j\in\mathcal{N}_i} & \mathrm{Col}[C_j(C_i^TC_i)^{-1}C_i^T]_{j\in\mathcal{N}_i}\mathcal{R}[Q_iC_i(C_i^TC_i)^{-1}C_j^T]_{j\in\mathcal{N}_i} & g_i\mathrm{Col}[C_j(C_i^TC_i)^{-1}C_i^TQ_iC_Z^T]_{j\in\mathcal{N}_i} \\ -g_i(d_i+g_i)C_Z(C_i^TC_i)^{-1}C_i^TQ_i & g_i\mathcal{R}[C_ZQ_iC_i(C_i^TC_i)^{-1}C_j^T]_{j\in\mathcal{N}_i} & g_i^2C_Z(C_i^TC_i)^{-1}C_i^TQ_iC_i(C_i^TC_i)^{-1}C_Z^T \end{bmatrix} \tag{5}$$

where $\rho_i \in \mathbb{R}^{4+2|\mathcal{N}_i|}$ is the costate variable. For each $i \in \mathcal{N}$, the left-hand side of the optimizations in (7) can be viewed as a value function, $V_i^*(\tilde{x}_i(0)) := \min_{u_i} J_i(\tilde{x}_i(0); u_i, u_{\mathcal{N}_i}^*), \forall i \in \mathcal{N}$ that should satisfy the appropriate coupled Hamilton-Jacobi equations,

$$H_i(\tilde{x}_i, \frac{\partial V_i^*}{\partial \tilde{x}_i}, u_i^*, u_{\mathcal{N}_i}^*) = 0,\ \forall i \in \mathcal{N}, \tag{8}$$

with boundary conditions $V_i^*(0) = 0$ and control policies given by, $u_i^* = -\theta\{\frac{1}{2}\begin{bmatrix}\frac{1}{L_i} \\ 0_{3+2|\mathcal{N}_i|}\end{bmatrix}^T \frac{\partial V_i^*}{\partial \tilde{x}_i}\}$.

A. Integral Reinforcement Learning

The formulation of the local distributed coupled Bellman equations (8) with $\tilde{Q}$ picked as in (5) to achieve tracking works only for finite horizon optimization. The problem is that (4) will have an infinite value and so that the optimization (6) is not well defined. Moreover the Bellman equations (8) require full knowledge of the dynamics of each VSI which in real world problems will limit the application of the proposed algorithm because of unknown quantities and unmodeled dynamics. From [21] we know that an equivalent formulation of the coupled Bellman equations (8) that does not involve the dynamics and produces a finite cost is given as,

$$\begin{aligned} V_i(\tilde{x}(t-T)) = \int_{t-T}^{t} \big(R_s(u_i^*) + \sum_{j\in\mathcal{N}_i} R_s(u_j^*) \\ + \tilde{x}_i^T\tilde{Q}_i\tilde{x}_i\big)d\tau + V_i(\tilde{x}(t)),\ \forall i \in \mathcal{N}, \end{aligned} \tag{9}$$

for any time $t \geqslant 0$ and time interval $T > 0$.

B. Existence and Stability of Equilibria

The next theorem provides a sufficient condition on the existence of a Nash equilibrium based on a relaxed version of the coupled Hamilton-Jacobi equations (9), where the equality to zero was replaced by less than or equal to zero.

Theorem 1: Suppose that the communication graph between the inverters is strongly connected and that there exist positive definite functions $V_i \in \mathcal{C}^1,\ i \in \mathcal{N}$ that satisfy the following inequalities,

$$\begin{aligned} &\int_{t-T}^{t} \big(\dot{V}_i(\tilde{x}_i) + R_s(u_i^*) \\ &+ \sum_{j\in\mathcal{N}_i} R_s(u_j^*) + \tilde{x}_i^T\tilde{Q}_i\tilde{x}_i\big)d\tau \leqslant 0, \forall t,\ \forall \tilde{x}_i, i \in \mathcal{N}, \end{aligned}$$

with $T \in \mathbb{R}^+$, $V_i(0) = 0,\ \forall i \in \mathcal{N}$ and also that $\tilde{Q}_i,\ \forall i \in \mathcal{N}$ is picked according to (5). The closed-loop systems (3) $\forall i \in \mathcal{N}$ with control policies

$$u_i^* = -\theta\{\frac{1}{2}\begin{bmatrix}\frac{1}{L_i} \\ 0_{3+2|\mathcal{N}_i|}\end{bmatrix}^T \frac{\partial V_i}{\partial \tilde{x}_i}\},\ \forall i \in \mathcal{N} \tag{10}$$

have asymptotically stable equilibrium points, and the inverters' outputs $y_i,\ \forall i \in \mathcal{N}$ are will all asymptotically synchronize to the exosystem. Moreover the control policies (10) form a Nash equilibrium and $J_i(\tilde{x}_i(0); u_i^*, u_{\mathcal{N}_i}^*) = V_i^*(\tilde{x}_i(0)),\ \forall i \in \mathcal{N}$.

Proof. The proof follows from [21] and is ommitted due to space limitations.

IV. Synchronization Using Actor/Critic Structures

Since in general it is not easy to solve (9) to find the costs $V_i,\ \forall i \in \mathcal{N}$ we shall use an actor/critic network architecture for every inverter to provide approximations of its cost and controller. Specifically, N critic approximators will provide approximations of the value functions $V_i, i \in \mathcal{N}$ and N actor approximators will provide approximations of the controllers $u_i,\ i \in \mathcal{N}$ inside the interval $[-V_{\mathrm{idc}}, V_{\mathrm{idc}}],\ i \in \mathcal{N}$. The critic approximators have the following form,

$$V_i^*(\tilde{x}_i) = W_i^T\phi_i(\tilde{x}_i) + \epsilon_i(\tilde{x}_i),\ \forall \tilde{x}_i,\ i \in \mathcal{N} \tag{11}$$

where $\phi_i(\tilde{x}_i) : \mathbb{R}^{4+2|\mathcal{N}_i|} \to \mathbb{R}^{h_1}$ are smooth basis functions with h_1 linearly independent basis bounded by $\|\phi_i\| \leqslant \phi_{\mathrm{imax}}$ and $\left\|\frac{\partial \phi_i}{\partial \tilde{x}_i}\right\| \leqslant \phi_{\mathrm{idmax}}$, $W_i \in \mathbb{R}^{h_1}$ are the ideal weights bounded by $\|W_i\| \leqslant W_{\mathrm{imax}}$ and $\epsilon_i(\tilde{x}_i)$ denotes the corresponding residual errors with $\sup_{\tilde{x}_i}\|\epsilon_i\| \leqslant \epsilon_{\mathrm{imax}}$ and $\sup_{\tilde{x}_i}\left\|\frac{\partial \epsilon_i}{\partial \tilde{x}_i}\right\| \leqslant \epsilon_{\mathrm{idmax}}$. The optimal controllers are approximated by the following actor approximators, $u_i^* = {W_{iu}}^T\phi_{iu}(\tilde{x}_i) + \epsilon_{iu}(\tilde{x}_i),\ \forall \tilde{x}_i,\ i \in \mathcal{N}$ where $\phi_{iu}(\tilde{x}_i) : \mathbb{R}^{4+2|\mathcal{N}_i|} \to \mathbb{R}^{h_2}$ are smooth basis functions with h_2 linearly independent basis bounded as $\|\phi_{iu}\| \leqslant \phi_{\mathrm{iumax}}$, $\epsilon_{iu}(\tilde{x}_i) \in \mathbb{R}$ are the residual errors, and $W_{iu} \in \mathbb{R}^{h_2}$ are the optimal weights.

By using (11) in (9) one obtains,

$$\begin{aligned} W_i^T\phi_i(\tilde{x}_i(t)) - W_i^T\phi_i(\tilde{x}_i(t-T)) + \int_{t-T}^{t} \big(R_s(u_i^*) \\ + \sum_{j\in\mathcal{N}_i} R_s(u_j^*) + \tilde{x}_i^T\tilde{Q}_i\tilde{x}_i\big)d\tau = \epsilon_{\mathrm{i}\pi}, i \in \mathcal{N}, \end{aligned} \tag{12}$$

where $\epsilon_{\mathrm{i}\pi}$ is a function of the approximation error ϵ_i and $T > 0$. We define the integral reinforcement signal $\forall t \geqslant 0$ for every VSI as, $\pi_i(t) = \int_{t-T}^{t}\big(R_s(u_i^*) + \sum_{j\in\mathcal{N}_i} R_s(u_j^*) + \tilde{x}_i^T\tilde{Q}_i\tilde{x}_i\big)d\tau,\ i \in \mathcal{N}$, which allows us to re-write (12) as,

$$W_i^T\Delta\phi_i(\tilde{x}_i(t)) + \pi_i(t) = \epsilon_{\mathrm{i}\pi}, i \in \mathcal{N}, \tag{13}$$

where $\Delta\phi_i(\tilde{x}_i(t)) := \phi_i(\tilde{x}_i(t)) - \phi_i(\tilde{x}_i(t-T))$.

The approximate value function and the optimal policy for every VSI using current estimates $\hat{W}_i$ and $\hat{W}_{iu}$ of the ideal weights W_i and W_{iu} respectively are given by the following critic and actor approximators,

$$\hat{V}_i(\tilde{x}(t)) = \hat{W}_i^T \phi_i(\tilde{x}_i(t)),\ \forall \tilde{x}_i, \tag{14}$$

$$\hat{u}_i(\tilde{x}(t)) = \hat{W}_{iu}^T \phi_{iu}(\tilde{x}_i(t)),\ \forall \tilde{x}_i. \tag{15}$$

Towards making (13) hold with a zero error $\epsilon_{i\pi}$, our goal is to tune the weights $\hat{W}_i$ so that the following error is driven to zero, $e_i := \hat{W}_i^T \Delta\phi_i(\tilde{x}_i(t)) - \hat{\pi}_i,\ \forall \tilde{x}_i$ where $\hat{\pi}_i := \int_{t-T}^{t} \big(R_s(\hat{u}_i^*) + \sum_{j\in\mathcal{N}_i} R_s(\hat{u}_j^*) + \tilde{x}_i^T \tilde{Q}_i \tilde{x}_i\big) d\tau$.

For that reason, we will use a gradient descent to select the critic network weights $\hat{W}_i$ to minimize the squared error $E_i = \frac{1}{2}e_i^2$,

$$\dot{\hat{W}}_i = -\alpha_i \frac{\Delta\phi_i(\tilde{x}_i(t))\big(\Delta\phi_i(\tilde{x}_i(t))^T \hat{W}_i + \hat{\pi}_i\big)}{(\Delta\phi_i(\tilde{x}_i(t))^T \Delta\phi_i(\tilde{x}_i(t)) + 1)^2}, \tag{16}$$

where $\alpha_i \in \mathbb{R}^+$ determines the speed of convergence.

Defining the critic estimation error as $\tilde{W}_i := W_i - \hat{W}_i$, the critic estimation error dynamics are given by

$$\begin{aligned}\dot{\tilde{W}}_i = &-\alpha_i \frac{\Delta\phi_i(\tilde{x}_i(t))}{(\Delta\phi_i(\tilde{x}_i(t))^T \Delta\phi_i(\tilde{x}_i(t)) + 1)^2} \Delta\phi_i(\tilde{x}_i(t))^T \tilde{W}_i \\ &+ \alpha_i \frac{\Delta\phi_i(\tilde{x}_i(t))}{(\Delta\phi_i(\tilde{x}_i(t))^T \Delta\phi_i(\tilde{x}_i(t)) + 1)^2} \epsilon_{\mathrm{i}\pi} := -N_{\mathrm{iom}} + p_{\mathrm{ier}},\end{aligned} \tag{17}$$

where N_{iom} is the nominal system and p_{ier} is the perturbation due to error $\epsilon_{\mathrm{i}\pi}$.

Theorem 2: Let the tuning of the critic network for every VSI be given by (16). Then the nominal system defined by (17) with $p_{\mathrm{ier}} = 0$ is exponentially stable with its trajectories satisfying $||\tilde{W}_i(t)|| \leqslant ||\tilde{W}_i(t_0)||\kappa_{i1} e^{-\kappa_{i2}(t-t_0)}$, for $t > t_0 \geqslant 0$, for some $k_{i1}, k_{i2} \in \mathbb{R}^+$ provided that the signal $M_i := \frac{\Delta\phi_i(\tilde{x}_i(t))}{(\Delta\phi_i(\tilde{x}_i(t))^T \Delta\phi_i(\tilde{x}_i(t)) + 1)}$ is persistently exciting (PE) over the interval $[t, t+T_{\mathrm{PE}}]$ with $\int_t^{t+T_{\mathrm{PE}}} M_i M_i^T d\tau \geqslant \beta_i I,\ \beta_i \in \mathbb{R}^+$.

Proof. The proof follows from [21].

Now we have to define the error for every actor network $e_{iu} \in \mathbb{R}$ that we have to drive to zero. The error is constructed by using (10) with $\hat{V}_i$ from (14) and has the following form, $e_{iu} := \hat{W}_{iu}^T \phi_{iu}(\tilde{x}_i(t)) + \theta\{\frac{1}{2} \begin{bmatrix} \frac{1}{L_i} \\ 0_{3+2|\mathcal{N}_i|} \end{bmatrix}^T \frac{\partial\phi_i}{\partial\tilde{x}_i}^T \hat{W}_i\}$.

In order to drive e_{iu} to zero for every VSI $i \in \mathcal{N}$ we have to select $\hat{W}_{iu}$ such that the following criterion is minimized,

$$E_{iu} = \frac{1}{2} e_{iu}^T e_{iu},\ i \in \mathcal{N}. \tag{18}$$

Hence, the tuning for the actor network can be found by applying gradient descent to (18) as follows,

$$\begin{aligned}\dot{\hat{W}}_{iu} &= -\alpha_{iu} \frac{\partial E_{iu}}{\partial \hat{W}_{iu}} = -\alpha_{iu}\phi_{iu}(\tilde{x}_i(t)) e_{iu} \\ &= -\alpha_{iu}\phi_{iu}(\tilde{x}_i(t)) \Bigg(\hat{W}_{iu}^T \phi_{iu}(\tilde{x}_i(t)) \\ &+ \theta\{\frac{1}{2} \begin{bmatrix} \frac{1}{L_i} \\ 0_{3+2|\mathcal{N}_i|} \end{bmatrix}^T \frac{\partial\phi_i}{\partial\tilde{x}_i}^T \hat{W}_i\} \Bigg)^T,\ i \in \mathcal{N}\end{aligned} \tag{19}$$

where $\alpha_{iu} \in \mathbb{R}^+$ determines the speed of convergence.

By defining the actor estimation error as $\tilde{W}_{iu} := W_{iu} - \hat{W}_{iu}$ its error dynamics can be written as,

$$\begin{aligned}\dot{\tilde{W}}_{iu} = &-\alpha_{iu}\phi_{iu}(\tilde{x}_i(t))\phi_{iu}(\tilde{x}_i(t))^T \tilde{W}_{iu} \\ &- \alpha_{iu}\phi_{iu}(\tilde{x}_i(t))\epsilon_{iu} \\ &- \alpha_{iu}\phi_{iu}(\tilde{x}_i(t))\theta\{\frac{1}{2}\tilde{W}_i^T \frac{\partial\phi_i(\tilde{x}_i)}{\partial\tilde{x}_i} \begin{bmatrix} \frac{1}{L_i} \\ 0_{3+2|\mathcal{N}_i|} \end{bmatrix}\} \\ &- \alpha_{iu}\phi_{iu}(\tilde{x}_i(t))\theta\{\frac{1}{2}\frac{\partial\epsilon_i}{\partial\tilde{x}_i} \begin{bmatrix} \frac{1}{L_i} \\ 0_{3+2|\mathcal{N}_i|} \end{bmatrix}\}.\end{aligned} \tag{20}$$

V. Equal Current Sharing and User Defined Matrices in the Performance

The designer needs to make a trade-off to balance the performance between good exosystem tracking, in terms of voltage magnitude and frequency, and good current sharing control. This section provides guidelines for selecting the weighting matrices $\tilde{Q}_i$ in (4) to promote equal current sharing among the VSIs. To accomplish this we rewrite $\tilde{x}_i^T \tilde{Q}_i \tilde{x}_i$ in (4) as follows, $\tilde{x}_i^T \tilde{Q}_i \tilde{x}_i := f_i + \big(\cdots\big)$, with $f_i := \big[i_i\ \mathcal{R}_{[i_j]_{j\in\mathcal{N}_i}}\big] Q_{fi} \big[i_i\ \mathcal{R}_{[i_j]_{j\in\mathcal{N}_i}}\big]^T$ and $\big(\cdots\big)$ being the terms of (4) that depend on the voltages and $Q_{fi} \in \mathbb{R}^{(|\mathcal{N}_i|+1)\times(|\mathcal{N}_i|+1)}$ has the diagonal form $Q_{fi} := \begin{bmatrix} Q_{fi11} & 0 \\ 0 & Q_{fi22} \end{bmatrix} \succ 0$, where the positive scalar $Q_{fi11} \in \mathbb{R}^+$ will weight the current of the i-th VSI and the diagonal matrix with positive entries $Q_{fi22} \in \mathbb{R}^{|\mathcal{N}_i|\times|\mathcal{N}_i|}$ will weight the currents in the neighborhood. The following lemma examines how picking appropriate user defined matrices matrices $\hat{Q}_i$ in (2) achieves equal current sharing among all inverters.

Lemma 1: Consider the following optimization of all current component terms i_i in $\tilde{x}_i^T \tilde{Q}_i \tilde{x}_i$:

$$\begin{aligned} &\text{minimize} && \sum_{i=1}^{N} f_i \\ &\text{subject to} && \sum_{i=1}^{N} i_i = \frac{V_{Ci}}{Z_{\mathrm{load}}} + \frac{dV_{Ci}}{dt}\sum_{i=1}^{N} C_{\mathrm{ap}_i} = i \\ &\text{w.r.t.} && i_i. \end{aligned}$$

Picking the weighting as,

$$Q_{fi11} + k_i = 1, \tag{21}$$

where $k_i \in \mathbb{R}^+$ is the weighting of the current i_i that appears in the neighborhood of other inverters, the minimum to this optimization is achieved at the equal current sharing $i_i = \frac{i}{N}$.

Proof. In order to minimize $F := \sum_{i=1}^{N} f_i$ subject to the constraint $\sum_{i=1}^{N} i_i = \frac{V_{Ci}}{Z_{\text{load}}} + \frac{dV_{Ci}}{dt}\sum_{i=1}^{N} C_{\text{ap}_i} = i$, in other words find the user-defined matrix Q_{fi} such that the current is equally shared, we apply the Lagrange multiplier method to write,

$$\begin{aligned} &\text{minimize} && F \\ &\text{subject to} && \sum_{i=1}^{N} i_i = i \\ &\text{w.r.t.} && i_i. \end{aligned}$$

Before we solve the optimization problem we need to expand,

$$\begin{aligned} F &= \sum_{i=1}^{N} \begin{bmatrix} i_i & \mathcal{R}_{[i_j]_{j\in\mathcal{N}_i}} \end{bmatrix} \begin{bmatrix} Q_{fi11} & 0 \\ 0 & Q_{fi22} \end{bmatrix} \begin{bmatrix} i_i & \mathcal{R}_{[i_j]_{j\in\mathcal{N}_i}} \end{bmatrix}^T \\ &= \sum_{i=1}^{N} i_i^2 Q_{fi11} + \sum_{i=1}^{N} \mathcal{R}_{[i_j]_{j\in\mathcal{N}_i}} Q_{fi22} \mathcal{R}^T_{[i_j]_{j\in\mathcal{N}_i}} \\ &\equiv \sum_{i=1}^{N} \Big(i_i^2 Q_{fi11} + i_i^2 k_i + \big(\mathcal{R}_{[i_j]\in\mathcal{N}_i} Q_{fi22} \mathcal{R}^T_{[i_j]_{j\in\mathcal{N}_i}}\big)_{i\neq j} \Big). \end{aligned}$$

Moreover let, $g := \sum_{i=1}^{N} i_i - i$ and let λ be a Lagrange multiplier to write, $\frac{\partial(F+\lambda g)}{\partial i_i} = 0$ which gives,

$$2 i_i \big(Q_{fi11} + k_i\big) + \lambda = 0. \tag{22}$$

We shall take for all $i \in \mathcal{N}$, $\frac{\partial(F+\lambda g)}{\partial\lambda} = 0$, $\sum_{i=1}^{N} i_i - i = 0$. After summing up $2i_i\big(Q_{fi11} + k_i\big) + \lambda = 0$ over all VSIs one has,

$$\lambda = -\frac{2\sum_{i=1}^{N} i_i \big(Q_{fi11} + k_i\big)}{N}, \tag{23}$$

and after substituting (23) in (22), yields, $i_i = \frac{\sum_{i=1}^{N} i_i \left(Q_{fi11}+k_i\right)}{N\left(Q_{fi11}+k_i\right)}$ which after taking into consideration (21) can be written as, $i_i = \frac{i}{N}$, that is the current is shared equally.

VI. Convergence and Stability Analysis

Fact 1: The following bounds are used to simplify the results from the following Theorem and are consequences of previous Assumptions,

$$\|\theta(\cdot)\| \leqslant V_{\text{idc}};\ \|\epsilon_{\text{i}\pi}\| \leqslant \epsilon_{\text{i}\pi\text{max}},\ \forall i \in \mathcal{N}.$$

□

To deal with the presence of the approximation errors with known bounds, and to obtain an asymptotically stable closed-loop system, one needs to add a robustifying term to the control (15) of the form,

$$\eta_i(t) = -\begin{bmatrix} \frac{1}{L_i} \\ 0_{3+2|\mathcal{N}_i|} \end{bmatrix} \frac{\|\tilde{x}_i\|^2}{K_{i1} + \|\tilde{x}_i\|^2} K_{i2},\ \forall i \in \mathcal{N} \tag{24}$$

where $K_{i1}, K_{i2} \in \mathbb{R}^+$ satisfy $\forall i \in \mathcal{N}$,

$$\begin{aligned} K_{i2} > L_i & \frac{K_{i1} + \|\tilde{x}_i\|^2}{\|\tilde{x}_i\|^2 \left(W_{\text{imax}}\phi_{\text{idmax}} + \epsilon_{\text{idmax}}\right)} \\ & \Big\{ \frac{\epsilon^2_{\text{iumax}}}{L_i^2} + 2\big(W_{\text{imax}}\phi_{\text{idmax}} + \epsilon_{\text{idmax}}\big)^2 + \big(\frac{1}{2}\epsilon_{\text{i}\pi\text{max}}\big)^2 \\ & + \Big(\phi_{\text{iumax}}\epsilon_{\text{iumax}} + 2\phi_{\text{iumax}} V_{\text{idc}}\Big)^2 \Big\}. \end{aligned} \tag{25}$$

With these terms the system dynamics (3), (15) by using the robustifying term (24) can be written as,

$$\dot{\tilde{x}}_i = \text{diag}\left(A_i, [A_j]_{j\in\mathcal{N}_i}, g_i A_Z\right)\tilde{x}_i + \begin{bmatrix} X_i \\ \text{Col}\big[X_j\big]_{j\in\mathcal{N}_i} \\ 0_2 \end{bmatrix}, \tag{26}$$

where,

$$X_i := \begin{bmatrix} \frac{1}{L_i} \\ 0 \end{bmatrix} \big((W^*_{iu} - \tilde{W}_{iu})^T \phi_{iu}(\tilde{x}_i) - K_{i2} \frac{\|\tilde{x}_i\|^2}{K_{i1} + \|\tilde{x}_i\|^2} \big).$$

A pseudocode that describes the algorithm has the following form,

Algorithm 1: Optimal Parallel Operation of VSIs

1: **procedure**
2: Start with initial currents and voltages in the neighboring augmented state $\tilde{x}_i(0)$, random initial weights $\hat{W}_i(0), \hat{W}_{iu}(0)$ for every inverter and construct the weighting $\tilde{Q}_i$ according to (21).
3: Propagate $t, \tilde{x}_i(t)$.
4: Propagate $\hat{W}_i(t), \hat{W}_{iu}(t)$ ▷ $\dot{\hat{W}}_{iu}$ as in (19) and $\dot{\hat{W}}_i$ as in (16).
5: Compute (14), and (15) for every inverter.
6: **end procedure**

The following theorem is the main result and proves asymptotic stability of the overall system.

Theorem 3: Consider the dynamics (3) for each VSI with the control input given by (15) and the $\tilde{Q}_i$, $\forall i \in \mathcal{N}$ picked according to (5). Assume that the signal $\frac{\Delta\phi_i(\tilde{x}_i(t))}{(\Delta\phi_i(\tilde{x}_i(t))^T \Delta\phi_i(\tilde{x}_i(t))+1)}$ is PE. The tuning laws for the critic and actor networks are given by (16) and (19) respectively and the learning information collecting interval T is picked sufficiently small. Then the equilibrium point of $\big(\tilde{x}_i(t), \tilde{W}_i(t), \tilde{W}_{iu}(t)\big)$, $\forall i \in \mathcal{N}$ is asymptotically stable for all $\big(\tilde{x}_i(0), \tilde{W}_i(0), \tilde{W}_{iu}(0)\big)$, $\forall i \in \mathcal{N}$ and the outputs of the VSIs at the load provide the appropriate voltage frequency and magnitude by following the exosystem, provided that the following inequalities are satisfied,

$$L_i \leqslant \frac{1}{\sqrt{2}} \approx 707.11mH;\ \alpha_i \geqslant \frac{1}{2},\ \forall i \in \mathcal{N}. \tag{27}$$

Proof. Consider the following Lyapunov equation $\mathcal{V}$ with $\mathcal{V}_i : \mathbb{R}^{(4+2|N_i|)} \times \mathbb{R}^{N_1} \times \mathbb{R}^{N_2} \to \mathbb{R}$ for all $t \geqslant 0$ defined as,

$$\begin{aligned}\mathcal{V} \equiv \sum_{1}^{N} \mathcal{V}_i &= \sum_{i=1}^{N} \left(V_i^* + V_{ic} + V_{iu}\right)\\ &= \sum_{i=1}^{N} \left(V_i^* + V_{ic} + \frac{1}{2\alpha_{iu}}\{\tilde{W}_{iu}^T \tilde{W}_{iu}\}\right)\end{aligned} \tag{28}$$

where V_i^* is the optimal value function for each inverter and $V_{ic} := \left\|\tilde{W}_i\right\|^2$ is a Lyapunov function for the critic error dynamics given by (17).

By taking the time derivative of (28) (first term with respect to (26), second term with respect to the perturbed critic estimation error dynamics (17), and substitute the actor error dynamics from (20)) and grouping the terms together, one obtains,

$$\begin{aligned}\dot{\mathcal{V}} = \sum_{i=1}^{N} \Bigg\{ &\frac{\partial V_i^*}{\partial \tilde{x}_i}^T \Bigg(\operatorname{diag}\left(A_i, [A_j]_{j\in\mathcal{N}_i}, g_i A_Z\right)\tilde{x}_i\\ &+ \begin{bmatrix} X_i \\ \mathrm{Col}\big[X_j\big]_{j\in\mathcal{N}_i} \\ 0_2 \end{bmatrix} \Bigg)\\ &- \frac{\partial V_{ic}}{\partial \tilde{W}_i}^T \left(\frac{\Delta\phi_i(\tilde{x}_i(t))\Delta\phi_i(\tilde{x}_i(t))^T}{(\Delta\phi_i(\tilde{x}_i(t))^T\Delta\phi_i(\tilde{x}_i(t))+1)^2} \right)\tilde{W}_i\\ &+ \frac{\partial V_{ic}}{\partial \tilde{W}_i}^T \Big(\frac{\Delta\phi_i(\tilde{x}_i(t))}{(\Delta\phi_i(\tilde{x}_i(t))^T\Delta\phi_i(\tilde{x}_i(t))+1)^2} \epsilon_{\mathrm{i}\pi} \Big)\\ &+ \tilde{W}_{iu}^T \Bigg(-\phi_{iu}(\tilde{x}_i(t))\phi_{iu}(\tilde{x}_i(t))^T \tilde{W}_{iu}\\ &- \phi_{iu}(\tilde{x}_i(t))\epsilon_{iu} - \phi_{iu}(\tilde{x}_i(t))\theta\{\frac{1}{2}\tilde{W}_i^T \frac{\partial\phi_i(\tilde{x}_i)}{\partial \tilde{x}_i} \begin{bmatrix} \frac{1}{L_i} \\ 0_{3+2|\mathcal{N}_i|} \end{bmatrix}\}\\ &- \phi_{iu}(\tilde{x}_i(t))\theta\{\frac{1}{2}\frac{\partial \epsilon_i}{\partial \tilde{x}_i} \begin{bmatrix} \frac{1}{L_i} \\ 0_{3+2|\mathcal{N}_i|} \end{bmatrix}\}\Bigg)\Bigg\}, t \geqslant 0.\end{aligned} \tag{29}$$

Now for simplicity we will take the following two terms of (29) separately as,

$$\begin{aligned}T_1 := \sum_{i=1}^{N} \Bigg\{ &\frac{\partial V_i^*}{\partial \tilde{x}_i}^T \Bigg(\operatorname{diag}\left(A_i, [A_j]_{j\in\mathcal{N}_i}, g_i A_Z\right)\tilde{x}_i\\ &+ \begin{bmatrix} X_i \\ \mathrm{Col}\big[X_j\big]_{j\in\mathcal{N}_i} \\ 0_2 \end{bmatrix} \Bigg)\Bigg\},\end{aligned} \tag{30}$$

and

$$\begin{aligned}T_2 := \sum_{i=1}^{N} \Bigg(&- \frac{\partial V_{ic}}{\partial \tilde{W}_i}^T \Big(\frac{\Delta\phi_i(\tilde{x}_i(t))\Delta\phi_i(\tilde{x}_i(t))^T}{(\Delta\phi_i(\tilde{x}_i(t))^T\Delta\phi_i(\tilde{x}_i(t))+1)^2} \Big)\tilde{W}_i\\ &+ \frac{\partial V_{ic}}{\partial \tilde{W}_i}^T \Big(\frac{\Delta\phi_i(\tilde{x}_i(t))}{(\Delta\phi_i(\tilde{x}_i(t))^T\Delta\phi_i(\tilde{x}_i(t))+1)^2} \epsilon_{\mathrm{i}\pi} \Big)\\ &+ \tilde{W}_{iu}^T \big(-\phi_{iu}(\tilde{x}_i(t))\phi_{iu}(\tilde{x}_i(t))^T \tilde{W}_{iu}\\ &- \phi_{iu}(\tilde{x}_i(t))\epsilon_{iu} - \phi_{iu}(\tilde{x}_i(t))\theta\{\frac{1}{2}\tilde{W}_i^T \frac{\partial\phi_i(\tilde{x}_i)}{\partial \tilde{x}_i} \begin{bmatrix} \frac{1}{L_i} \\ 0_{3+2|\mathcal{N}_i|} \end{bmatrix}\}\\ &- \phi_{iu}(\tilde{x}_i(t))\theta\{\frac{1}{2}\frac{\partial \epsilon_i}{\partial \tilde{x}_i} \begin{bmatrix} \frac{1}{L_i} \\ 0_{3+2|\mathcal{N}_i|} \end{bmatrix}\}\big)\Bigg).\end{aligned} \tag{31}$$

After using the Bellman equation (8) in (30) one obtains,

$$\begin{aligned}T_1 = \sum_{i=1}^{N} \Bigg(&- R_s(u_i^*) - \sum_{j\in\mathcal{N}_i} R_s(u_j^*) - \tilde{x}_i^T \tilde{Q}_i \tilde{x}_i\\ &- \frac{\partial V_i^*}{\partial \tilde{x}_i}^T \begin{bmatrix} \begin{bmatrix} \frac{1}{L_i} \\ 0 \end{bmatrix} \tilde{W}_{iu}^T \phi_{iu}(\tilde{x}_i) \\ \mathrm{Col}\Big[\begin{bmatrix} \frac{1}{L_j} \\ 0 \end{bmatrix} \tilde{W}_{ju}^T \phi_{ju}(\tilde{x}_j)\Big]_{j\in\mathcal{N}_i} \\ 0_2 \end{bmatrix}\\ &- \frac{\partial V_i^*}{\partial \tilde{x}_i}^T \begin{bmatrix} \begin{bmatrix} \frac{1}{L_i} \\ 0 \end{bmatrix} \epsilon_{iu} \\ \mathrm{Col}\Big[\begin{bmatrix} \frac{1}{L_j} \\ 0 \end{bmatrix} \epsilon_{ju}\Big]_{j\in\mathcal{N}_i} \\ 0_2 \end{bmatrix}\\ &- \frac{\partial V_i^*}{\partial \tilde{x}_i}^T \begin{bmatrix} \begin{bmatrix} \frac{1}{L_i} \\ 0 \end{bmatrix} K_{i2}\frac{\|\tilde{x}_i\|^2}{K_{i1}+\|\tilde{x}_i\|^2} \\ \mathrm{Col}\Big[\begin{bmatrix} \frac{1}{L_j} \\ 0 \end{bmatrix} K_{j2}\frac{\|\tilde{x}_j\|^2}{K_{j1}+\|\tilde{x}_j\|^2}\Big]_{j\in\mathcal{N}_i} \\ 0_2 \end{bmatrix}\Bigg)\end{aligned}$$

which can be upper bounded by using vector nested norms as,

$$\begin{aligned}T_1 \leqslant \sum_{i=1}^{N} \Bigg(&- R_s(u_i^*) - \sum_{j\in\mathcal{N}_i} R_s(u_j^*) - \tilde{x}_i^T \tilde{Q}_i \tilde{x}_i\\ &+ (W_{\mathrm{imax}}\phi_{\mathrm{idmax}} + \epsilon_{\mathrm{idmax}})\\ &\Big(\sqrt{\frac{\phi_{\mathrm{iumax}}^2}{L_i^2}\left\|\tilde{W}_{iu}\right\|^2 + \sum_{j\in\mathcal{N}_i}\frac{\phi_{\mathrm{jumax}}^2}{L_j^2}\left\|\tilde{W}_{ju}\right\|^2}\\ &+ \sqrt{\frac{\epsilon_{\mathrm{iumax}}^2}{L_i^2} + \sum_{j\in\mathcal{N}_i}\frac{\epsilon_{\mathrm{jumax}}^2}{L_j^2}}\\ &- \Big(\left\|\frac{K_{i2}}{L_i}\frac{\|\tilde{x}_i\|^2}{K_{i1}+\|\tilde{x}_i\|^2}\right\|_1 + \sum_{j\in\mathcal{N}_i}\left\|\frac{K_{j2}}{L_j}\frac{\|\tilde{x}_j\|^2}{K_{j1}+\|\tilde{x}_j\|^2}\right\|_1\Big)\Big)\Bigg).\end{aligned} \tag{32}$$

After using Young's inequality for products, one can further upper bound (32) as,

$$\begin{aligned}T_1 \leqslant \sum_{i=1}^{N} \Bigg(&- R_s(u_i^*) - \sum_{j\in\mathcal{N}_i} R_s(u_j^*) - \underline{\lambda}(\tilde{Q}_i)\left\|\tilde{x}_i\right\|^2\\ &+ \left(W_{\mathrm{imax}}\phi_{\mathrm{idmax}} + \epsilon_{\mathrm{idmax}}\right)^2 + \frac{1}{2}\frac{\phi_{\mathrm{iumax}}^2}{L_i^2}\left\|\tilde{W}_{iu}\right\|^2\\ &+ \frac{1}{2}\sum_{j\in\mathcal{N}_i}\frac{\phi_{\mathrm{jumax}}^2}{L_j^2}\left\|\tilde{W}_{ju}\right\|^2 + \frac{1}{2}\frac{\epsilon_{\mathrm{iumax}}^2}{L_i^2} + \frac{1}{2}\sum_{j\in\mathcal{N}_i}\frac{\epsilon_{\mathrm{jumax}}^2}{L_j^2}\\ &- (W_{\mathrm{imax}}\phi_{\mathrm{idmax}} + \epsilon_{\mathrm{idmax}})\Big(\frac{K_{i2}}{L_i}\frac{\|\tilde{x}_i\|^2}{K_{i1}+\|\tilde{x}_i\|^2}\\ &+ \sum_{j\in\mathcal{N}_i}\frac{K_{j2}}{L_j}\frac{\|\tilde{x}_j\|^2}{K_{j1}+\|\tilde{x}_j\|^2}\Big)\Bigg).\end{aligned}$$

Now we can upper bound (31) as,

$$T_2 \leqslant \sum_{j\in\mathcal{N}_i}\Big(-\alpha_i\left\|\tilde{W}_i\right\|^2-\phi_{\text{iumax}}^2\left\|\tilde{W}_{iu}\right\|^2+\frac{1}{2}\epsilon_{\text{i}\pi\text{max}}\left\|\tilde{W}_i\right\| \\ -\big(\phi_{\text{iumax}}\epsilon_{\text{iumax}}+2\phi_{\text{iumax}}V_{\text{idc}}\big)\left\|\tilde{W}_{iu}\right\|\Big). \tag{33}$$

By using Young's inequality for products one can further upper bound (33) as,

$$T_2 \leqslant \sum_{j\in\mathcal{N}_i}\Big(-\alpha_i\left\|\tilde{W}_i\right\|^2-\phi_{\text{iumax}}^2\left\|\tilde{W}_{iu}\right\|^2 \\ +\frac{1}{2}\big(\frac{1}{2}\epsilon_{\text{i}\pi\text{max}}\big)^2+\frac{1}{2}\left\|\tilde{W}_i\right\|^2 \\ +\frac{1}{2}\Big(\big(\phi_{\text{iumax}}\epsilon_{\text{iumax}}+2\phi_{\text{iumax}}V_{\text{idc}}\big)\Big)^2+\frac{1}{2}\left\|\tilde{W}_{iu}\right\|^2\Big).$$

Since we are summing over all the inverters in the terms T_1 and T_2 and since the sum over the neighborhood has similar properties we can cross out the similar terms. Finally (29) can be written as, $\dot{\mathcal{V}} \leqslant T_1 + T_2$ and after taking into consideration the bound of K_{i2} given by (25) one has,

$$\dot{\mathcal{V}} \leqslant \sum_{i=1}^{N}\Big(-R_s(u_i^*)-\sum_{j\in\mathcal{N}_i}R_s(u_j^*)-\underline{\lambda}(\tilde{Q}_i)\left\|\tilde{x}_i\right\|^2 \\ -(\phi_{\text{iumax}}^2-\frac{1}{2}\frac{\phi_{\text{iumax}}^2}{L_i^2})\left\|\tilde{W}_{iu}\right\|^2-(\alpha_i-\frac{1}{2})\left\|\tilde{W}_i\right\|^2\Big).$$

By taking into account the inequalities (27) one has $\dot{\mathcal{V}} < 0, t \geqslant 0$. From Barbalat's lemma [11] it follows that as $t \to \infty$, then $\forall i \in \mathcal{N}$, $\|e_{\text{rr}i}\| \to 0$, $\left\|\tilde{W}_i\right\| \to 0$ and $\left\|\tilde{W}_{iu}\right\| \to 0$ and hence the robustifying term goes to zero asymptotically. Finally since the robustifying term goes to zero hence $\|\hat{u}_i\| \to u_i^*$.

VII. Implementation Issues and Simulations

A. Implementation Discussion

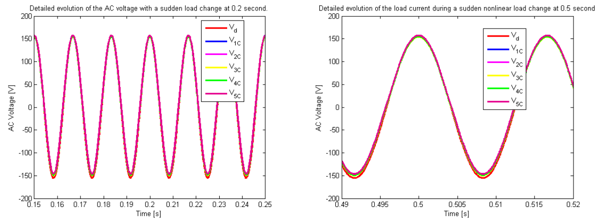
Since the IGBTs (Insulated Gate Bipolar Transistors) have become the component of choice for many power electronic applications, we need to open and close appropriate switches to take $\pm V_{\text{idc}}$. For that reason, the control input (15) has to provide DC voltage $\{-V_{\text{idc}}, V_{\text{idc}}\}$ and not $[-V_{\text{idc}}, V_{\text{idc}}]$. To do that, one should pass the control input through a hard limiter that restricts its output to one of the two values $\pm V_{\text{idc}}$, i.e.

$$\hat{u}_{\text{idc}} = V_{\text{idc}}\ \theta_{\text{hlim}}(\hat{u}_i),\ i \in \mathcal{N}$$

where $\forall i \in \mathcal{N},\ \theta_{\text{hlim}}(\hat{u}_i) = \begin{cases}-1, & \hat{u}_i < 0\\ 1, & \hat{u}_i \geqslant 0.\end{cases}$

B. Algorithmic Evaluation

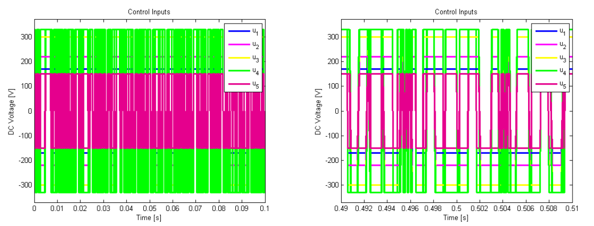
Consider the configuration of inverters shown in Figure 1 with $L_1 = 1.1mH, L_2 = 1.2mH, L_3 = 3mH, L_4 = 1.263mH, L_5 = 1.187mH, C_{\text{ap1}} = 41\mu F, C_{\text{ap2}} = 20\mu F, C_{\text{ap3}} = 50\mu F, C_{\text{ap4}} = 40.3\mu F, C_{\text{ap5}} = 39.6\mu F, R_1 = 0.1\Omega, R_2 = 0.15\Omega, R_3 = 0.1\Omega, R_4 = 0.2\Omega, R_5 = 0.15\Omega$. Since the output voltage fundamental AC frequency (60 Hz) is quite low the line impedances can be considered as resistive and we take $Z_1 = 0.01\Omega, Z_2 = 0.015\Omega, Z_3 = 0.02\Omega, Z_4 = 0.04\Omega, Z_5 = 0.035\Omega$. The basis functions for the critic are picked as quadratic in the neighboring state and for the actor as quadratic and linear in the neighboring state. the user defined matrices are defined by using Lemma 1, as $\hat{Q}_i = \begin{bmatrix}1 & 0\\ 0 & 100\end{bmatrix}$, $\forall i \in \mathcal{N}$ and $Q_{ri} = 100,\ \forall i \in \mathcal{N},\ \tilde{Q}_i, \in \mathcal{N}$ are found from (5) and the control inputs are $\pm V_{\text{idc}}$ with $V_{1\text{dc}} = 170\text{Volts},\ V_{2\text{dc}} = 220\text{Volts},\ V_{3\text{dc}} = 300\text{Volts},\ V_{4\text{dc}} = 350\text{Volts},\ V_{5\text{dc}} = 150\text{Volts}$ after going through the hard limiter. The outputs of the inverters are corrupted by unbiased band-limited white noise with variable strengths. We emphasize that the control algorithm does not require the values of the parasitic resistances $R_i, \forall i, \in \mathcal{N}$, the capacitances $C_{\text{api}},\ \forall i, \in \mathcal{N}$, the load and the frequency of the exosystem ω_0. The tuning gains are picked as $\alpha_i = 200, \forall i, \in \mathcal{N}$ and $\alpha_{iu} = 0.1,\ \forall i, \in \mathcal{N}$ and time interval $T = 0.01$ sec. The algorithm from Theorem 3 is applied to a case of no-load connected from $t < 0.2$, a linear load $Z_{\text{load}} = R_{\text{load}}$ (in the algorithm proposed the load is considered unknown) from $0.2 \leqslant t < 0.5$ and then a nonlinear load model adopted from a fluorescent lamp model with an unknown equivalent resistance variation [5] is added for $t \geqslant 0.5$. One can see that the algorithm is very robust to sudden load changes. In Figures 3(a)-3(b) one can see the desired output voltage versus the actual one. The sudden changes of load are shared amongst the inverters and their magnitudes and phases are strictly equal to each other. One can see in Figure 4 the control inputs (pulse-like waveform of $\pm V_{\text{idc}},\ \forall i \in \mathcal{N}$) applied to the VSIs. Figures 5 and 6 show the sharing of the load during changes at each inverter. To characterize the power quality of our VSIs, we use the total harmonic distortion (THD) which is the ratio of the sum of the powers of all harmonic components to the power of the fundamental harmonic. One should obtain low THDs ($< 3\%$) in order to have reduction in peak currents, namely the proposed framework provides THDs which are essentially below 1.3%, 1.42%, 1.65%, 1.6% and 1.38%. This indicates that the controllers can operate the system effectively in the presence of sudden load changes and white noise in the output. Finally, the fact that a hard limiter hardly makes any change in the performance confirms that the method is robust. This is not surprising in view of the fact that we use a game-theoretical framework that protects the system against worst-case disturbances.



(a) Evolution of the AC voltage during the linear load change.

(b) Evolution of the AC voltage during the nonlinear load change.

Fig. 3. AC Voltage with a magnitude of $110\sqrt{2}$ Volts and a frequency of 60 Hz is produced at the output.



(a) Control input provides the appropriate DC voltage.

(b) Zoomed control input during the sudden load change (from linear to nonlinear).

Fig. 4. Control input is $\pm V_{\rm idc}$.

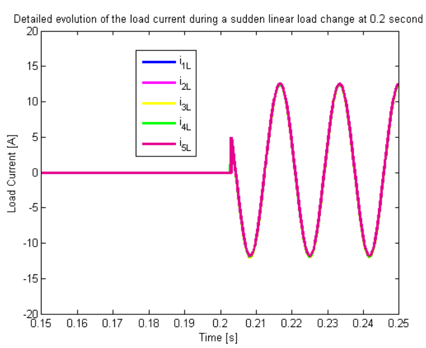


Fig. 5. Detailed evolution of the load currents during a sudden connection of a linear load at 0.2.

VIII. Conclusion and Future Work

This paper proposed a learning based optimal synchronization algorithm that provides the desired magnitude and frequency at the AC appliances by integrating separate DC/AC micro-inverters to convert DC voltage to AC voltage. The computational intelligent system creates a high quality, self-sufficient AC grid. The controllers ensure exosystem tracking, in terms of AC voltage magnitude and frequency, with guaranteed performance and equal current sharing. In order to implement this algorithm we design distributed optimal controllers that use current and voltage information in the $L-C$ filters from neighboring inverters and without requiring

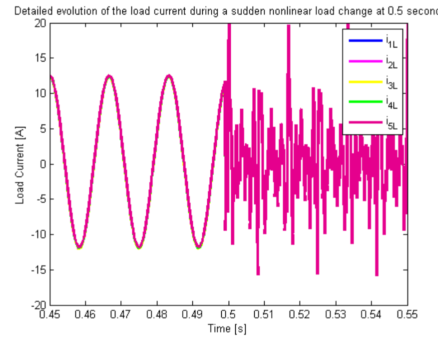


Fig. 6. Detailed evolution of the load currents during a sudden connection of a nonlinear load at 0.5.

the knowledge of the values of the load and of any parasitic resistances or capacitances which are hard to know. The algorithm builds upon the ideas of ADP and uses only partial information of the systems and the exosystem which is pinned to very few inverters in order to provide the desired behavior without phasor domain analysis. Simulation results showed the effectiveness of the proposed approach. Future work will be concentrated on testing the present algorithm in hardware and to three-phase systems with delays and disturbance attenuation mechanisms (i.e. $\mathcal{H}_\infty$).

References

[1] B. D. O Anderson and J. B. Moore, *Linear Optimal Control*, Englewood Cliffs, Prentice Hall, NJ, 1971.

[2] T. Basar, G. J. Olsder, *Dynamic noncooperative game theory* (2nd ed.), Philadelphia, PA: SIAM, 1999 .

[3] K. de Brabandere, B. Bolsens, J. Van den Keybus, A. Woyte, J. Driesen, R. Belmans, "A voltage and frequency droop control method for parallel inverters," *IEEE Trans. Power Electron.*, vol. 22, no. 4, pp. 1107-1115, July 2007.

[4] E. J. Bueno, F. Espinosa, F. J. Rodriguez, J. Urena, S. Cobreces. "Current Control of Voltage Source Converters connected to the grid through an LCL- filter," *Proc. IEEE Power Electronics Specialists Conference*, pp. 68-73, June 2004.

[5] M. Cervi, A. R. Seidel, F. E. Bisogno and D. Prado, "Fluorescent Lamp Model based on the Equivalent Resistance Variation," *Proc. IEEE Intl. Conf. on Ind. Appl.*, pp. 680-684, February 2002.

[6] F. Dorfler, M. R. Jovanovic, M. Chertkov, and F. Bullo, "Sparse and optimal wide-area damping control in power networks," *Proc. American Control Conference*, pp. 4295-4300, June 2013.

[7] J. M. Guerrero,J. Matas, Luis Garcia de Vicuna, M. Castilla, J. Miret, J., "Decentralized Control for Parallel Operation of Distributed Generation Inverters Using Resistive Output Impedance," *IEEE Trans. Ind. Electron.*, vol. 54, no. 2, pp. 994-1004, March 2007.

[8] R.M. Santos Filho, P. F. Seixas, P.C. Cortizo, L. Torres, A.F. Souza, "Comparison of Three Single-Phase PLL Algorithms for UPS Applications," *IEEE Trans. Ind. Electron.*, vol.55, no.8, pp. 2923-2932, September 2008.

[9] M. Hua, H. Hu, Y. Xing, Z. He, "Distributed control for AC-motor-drive-inverters in parallel operation," *IEEE Trans. Ind. Electron.*, vol. 58, no. 12, pp. 5361-5370, December 2011.

[10] T. Kawabata, T. Miyashita, Y. Yamamoto. "Dead Beat Control of Three Phase PWM Inverter," *IEEE Trans. on Power Electron.*, vol. 5, no. 1, pp. 21-28, January 1990.

[11] H. K. Khalil, *Nonlinear Systems*, Prentice Hall, 2002.

[12] P. Kundur, N. J. Balu, and M. G. Lauby, *Power System Stability and Control*, New York: McGraw-Hill, 1994.

[13] T.B. Lazzarin, G. A. T. Bauer, I. Barbi, "A Control Strategy for Parallel Operation of Single-Phase Voltage Source Inverters: Analysis, Design and Experimental Results," *IEEE Trans. on Ind. Electron.*, vol.60, no.6, pp. 2194-2204, June 2013.

[14] M. Lindgren, J. Svensson, "Control of a Voltage-source Converter Connected to the Grid through an LCL-filter - Application to Active Filtering," *Proc. IEEE Power Electron. Specialist Conference*, pp. 229-235, May 1998.

[15] S. Mariethoz, A. G. Beccuti, M. Morari, "Analysis and optimal current control of a voltage source inverter connected to the grid through an LCL filter," *IEEE Power Electron. Specialists Conference*, pp. 2132-2138, June 2008.

[16] R. R. Negenborn, *Multi-agent model predictive control with applications to power networks*, TU Delft, Delft University of Technology, 2007

[17] M. Prodanovic, T. C. Green. "Control and Filter Design of Three-Phase Inverters for High Power Quality Grid Connection," *IEEE Trans. on Power Electronics*, vol. 18, no. 1, pp. 373-380, February 2003.

[18] J. W. Simpson-Porco, F. Dorfler, F. Bullo, "Synchronization and power sharing for droop-controlled inverters in islanded microgrids," *Automatica*, vol. 49, no. 9, pp. 2603-2611, September 2013.

[19] L.A.B. Torres, J. P. Hespanha, J. Moehlis, "Synchronization of Identical Oscillators Coupled Through a Symmetric Network With Dynamics: A Constructive Approach With Applications to Parallel Operation of Inverters," *IEEE Trans. on Automatic Control*, vol. 60, no. 12, pp. 3226-3241, December 2015.

[20] E. Twining, D. G. Holmes, "Grid Current Regulation of a Three-Phase Voltage Source Inverter With an LCL Input Filter," *IEEE Trans. on Power Electronics*, vol. 18, no. 3, pp. 888-895, May 2003.

[21] D. Vrabie, K. G. Vamvoudakis, F. L. Lewis, *Optimal Adaptive Control and Differential Games by Reinforcement Learning Principles*, Control Engineering Series, IET Press, 2012.

[22] W. Yao, M. Chen, J. Matas, J. M. Guerrero, Q. Zhao-ming, "Design and analysis of the droop control method for parallel inverters considering the impact of the complex impedance on the power sharing," *IEEE Trans. Ind. Electron.*, vol. 58, no. 2, pp. 576-588, February 2011.

[23] F. Zare and J. A. Firouzjaee, "Hysteresis band current control for a single phase Z-source inverter with symmetrical and asymmetrical Z-network," *Proc. Power Convers. Conf.*, pp. 143-148, April 2007.

[24] D. N. Zmood, D. G. Holmes, G. H. Bode, "Frequency-Domain Analysis of Three-Phase Linear Current Regulator," *IEEE Trans. on Indust. Applications*, vol. 37, No. 2, pp. 601-610, August 2002.

Kyriakos G. Vamvoudakis (SM'15) received the Diploma in ECE from Technical University of Crete, Greece in 2006 with highest honors. After moving to USA, he studied at The University of Texas and he received his M.S. and Ph.D. in ECE in 2008 and 2011 respectively. During the period from 2012 to 2016 he was a project research scientist at the Center for Control, Dynamical Systems and Computation at the University of California, Santa Barbara. He is now an assistant professor in the Department of Aerospace and Ocean Engineering at Virginia Tech. His research interests include approximate dynamic programming, game theory, and optimal control. Recently, his research has focused on network security, smart grid and multi-agent optimization. Prof. Vamvoudakis is the recipient of several international awards including the 2016 International Neural Network Society Young Investigator Award, the Best Paper Award for Autonomous Vehicles at the 27th Army Science Conference in 2010, and the Best Researcher Award from the Automation and Robotics Research Institute in 2011.

João P. Hespanha (F'08) received the Licenciatura in ECE from the IST, Lisbon, Portugal in 1991 and the Ph.D. in electrical engineering and applied science from Yale University in 1998. From 1999 to 2001, he was Assistant Professor at the University of Southern California. He moved to the University of California, Santa Barbara in 2002, where he currently is a Professor with the Department of Electrical and Computer Engineering. Prof. Hespanha is the Chair of the Department of Electrical and Computer Engineering. His current research interests include hybrid and switched systems; multi-agent control systems; distributed control over communication networks; and network security. Prof. Hespanha is the recipient of the Yale University's Henry Prentiss Becton Graduate Prize for exceptional achievement in research in Engineering and Applied Science, a National Science Foundation CAREER Award, the 2005 Automatica Theory/Methodology best paper prize, the 2006 George S. Axelby Outstanding Paper Award, and the 2009 Ruberti Young Researcher Prize.