Information Sciences 296 (2015) 25–41

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Transcale control for a class of discrete stochastic systems based on wavelet packet decomposition

Lin Zhao *, Yingmin Jia

The Seventh Research Division, Beihang University (BUAA), Beijing 100191, PR China

ARTICLE INFO

Article history:
Received 8 July 2014
Received in revised form 13 October 2014
Accepted 16 October 2014
Available online 24 October 2014

Keywords:
Transcale control
LQG control
Tracking control
Multiscale system
Wavelet packet decomposition

ABSTRACT

In this paper, a wavelet packet decomposition (WPD) based real-time transcale linear-quadratic-Gaussian (LQG) tracking control algorithm is given for a class of discrete stochastic systems, which is developed with the state-space models of wavelet packet coefficients at the coarsest scale decomposition layer established by Haar WPD. The WPD-based transcale control algorithm provides a compromise between performance index and computational efficiency compared with the conventional LQG tracking control algorithm, and it is an improvement of the wavelet transform based algorithm previously proposed. Simulation results are presented to demonstrate the effectiveness of the proposed algorithm.

1. Introduction

The linear-quadratic-Gaussian (LQG) control is an optimal control of linear systems with quadratic cost function in the presence of Gaussian noises, which has been deeply researched and many results have been obtained [1,2,16,21,22,37,20]. For example, You et al. [37] discussed the quantised LQG control for linear stochastic systems; Arantes et al. [20] investigated the application of LQG control in spacecraft attitude systems. Output tracking problem is one of the most common and important issues in designing a control system, which has wide applications in dynamic processes in industry, economics, and biology [15,8,31,38,4,39,40,36,41,9,27,33,10]. The main objective of tracking control is to make the output of the model, via a controller, track the output of a given reference model as closely as possible. The LQG-based tracking control algorithms in [8,31,38,4] can be regarded as the classical single scale system control method since they focus on system represented at a single scale and design the control law at that scale. In practice, however, the system or output may contain multiscale (or equivalently, multiresolution) features. Even if they do not have multiscale features, more confidence can be obtained by using the multiscale processing algorithm, which can reduce the uncertainty and complexity of the problems [5]. Moreover, the speed up computation of control actions have been studied in some optimal control to guarantee the complex control schemes in real time [34,14]. In our initial work [45], a wavelet transform (WT)-based transcale LQG tracking control algorithm was proposed, which brought LQG control into a multiscale scenario and effectively improved the computation efficiency of conventional LQG algorithm by parallel processing.

Models that describe process behavior at several spatial and temporal scales are essential for many engineering tasks ranging from process analysis and design to operational monitoring, control and optimization [29]. Multiscale processing algorithm is an effective tool to represent the phenomena occurred at different scales, see [23,25,5,11,12,17,42,18,28,19,6,7,32,43,29,30,26]. The

* Corresponding author.
E-mail addresses: zhaolin1585@163.com (L. Zhao), ymjia@buaa.edu.cn (Y. Jia).

http://dx.doi.org/10.1016/j.ins.2014.10.039
0020-0255/

wavelet transform (WT), which is a tool for describing multiscale data structures, has been used to develop multiscale processing algorithm. For example, Zhao et al. [43] used WT to develop a robust transcale state estimation algorithm for multiresolution discrete-time systems, which could not only fuse multiresolution sensor information but also improve the computation efficiency and estimation accuracy; Stephanopoulos et al. [29,30] introduced an alternative philosophy to establish the multiscale models by Haar WT and develop the multiscale model predictive control algorithm, which could also improve the computation efficiency.

Most of these approaches in [43,29,30,26,45] rely on the multiscale decomposition to systems or measurements through WT, but sometimes it is not an optimal choice for discrete sequences compared with wavelet packet decomposition (WPD) [13]. WPD is a transformation in which the signal passes through more filters than WT. For J scales decomposition, the WPD produces 2^J sets of coefficients as opposed to $J+1$ sets of coefficients in WT. However, due to the downsampling process the overall number of coefficients is still the same without redundancy. From the point of view of signal analysis, the standard WT may not produce the best result, since it is limited to wavelet basis [35,13]. In fact, the wavelet basis based WT is only a special WPD. From the point of view of computation, the decomposition structure of WPD gives us more chances to improve the computation efficiency compared with that of WT. So it is necessary to further study the WPD-based algorithm for control and estimation of discrete stochastic systems, which can ensure the multiscale processing algorithm be better combined with conventional control algorithm.

Motivated by the above discussions, in this study, we introduce a WPD-based real-time transcale LQG tracking control algorithm for the output tracking of discrete stochastic systems. The term *transcale control* is defined in [44,45], which is applied to emphasize the fact that such a control strategy is capable of designing the control law at the finest scale but reflecting the desired performances at different coarser scales. The Haar WPD is employed to establish the state-space models of wavelet packet coefficients (WPC) at the coarsest scale decomposition layer. Based on these models, the real-time transcale LQG tracking control algorithm is proposed by three steps. The computational efficiency analysis is given. It can be proved that the control law designed using the proposed algorithm has the desired performance in the meaning of transcale control.

In summary, the contributions of this paper are as follows:

(1) The LQG tracking control algorithm is further developed in the multiscale framework. This WPD-based algorithm can improve the computation efficiency compared with the LQG tracking control algorithm.
(2) The WPD is firstly proposed into the multiscale control for discrete stochastic systems (compared with [29,30,45]).
(3) The transcale LQG tracking control algorithm in [45] is developed based on WT while the decomposition under WT is a special WPD, so the WT-based algorithm in [45] is only a special case of the WPD-based algorithm developed in this paper.
(4) From the point of the view of computation efficiency, the decomposition structure in this paper is better than that of others, so the WPD-based algorithm can effectively improve the computation efficiency compared with the WT-based algorithm in [45].

The rest of this paper is organized as follows. The control problem formulation is summarized in Section 2. A brief introduction about the Haar WPD of discrete sequences is given in Section 3. The WPD-based control scheme is presented in Section 4 numerical example of the derived control design is presented in Section 5. Conclusions are given in Section 6.

2. Problem formulation

Consider the following discrete stochastic system

$$x_{k+1} = Ax_k + Bu_k + Dw_k \tag{1}$$

$$y_k = Cx_k + v_k \tag{2}$$

$$z_k = C_d x_k, \quad k = 0, 1, \ldots, N \tag{3}$$

where $x_k \in \mathrm{R}^n$, $u_k \in \mathrm{R}^m$, $y_k \in \mathrm{R}^q$ and $z_k \in \mathrm{R}^s$ are the state, control input, measurement, and output vectors respectively. $A \in \mathrm{R}^{n\times n}$, $B \in \mathrm{R}^{n\times m}$, $D \in \mathrm{R}^{n\times p}$ are the system matrices, $C \in \mathrm{R}^{q\times n}$ is the observation matrix, $C_d \in \mathrm{R}^{s\times n}$ is the output matrix. The system noise $w_k \in \mathrm{R}^p$ and the measurement noise $v_k \in \mathrm{R}^q$ are uncorrelated white zero-mean Gaussian vectors with covariance matrices $UandV$ respectively. The statistic characteristics of initial x_0 are given as $E\{x_0\} = m_0 \in \mathrm{R}^n$, $E\{(x_0 - m_0)(x_0 - m_0)^T\} = \overline{P}_0 \in \mathrm{R}^{n\times n}$.

In this paper, the finite-horizon output tracking problem is considered, where the desired output is given by $\bar{z}_k \in \mathrm{R}^s$, the time horizon $N = M2^J - 1$ and M is a positive integer. Denote

$$e_k \triangleq \bar{z}_k - z_k \tag{4}$$

The LQG tracking control can solve the above problem, which is to find u_k such that the following cost function is minimized

$$J(u_k) \triangleq \frac{1}{2} E\left\{ e_N^T F e_N + \sum_{k=0}^{N-1} (e_k^T Q e_k + u_k^T R u_k) \right\} \tag{5}$$

$F, Q \in \mathrm{R}^{s\times s}$ are positive-semidefinite matrices, $R \in \mathrm{R}^{m\times m}$ is positive-definite matrix, u_k is unconstrained. The following separation principle guarantees the Kalman filter and the linear-quadratic tracking regulator can be designed and computed independently.

Lemma 1. *To guarantee the cost function (6) is minimized, if (A, C_d) is observable, then the control law is designed as*

$$u_k = -[R + B^T P_{k+1} B]^{-1} B^T [P_{k+1} A \hat{x}_k + g_{k+1}] \tag{6}$$

where the real, symmetric and positive-definite matrix $P_k \in \mathrm{R}^{n\times n}$ is the solution to the Riccati-type matrix difference equation

$$P_k = C_d^T Q C_d + A^T P_{k+1} A - A^T P_{k+1} B (R + B^T P_{k+1} B)^{-1} B^T P_{k+1} A \tag{7}$$
$$P_N = F \tag{8}$$

The additional vector $g_k \in \mathrm{R}^n$ is a feed-forward term, which is the solution to the linear vector difference equation

$$g_k = -C_d^T Q \bar{z}_k - A^T P_{k+1} B (R + B^T P_{k+1} B)^{-1} B^T g_{k+1} + A^T g_{k+1} \tag{9}$$
$$g_N = -C_d^T F \bar{z}_N \tag{10}$$

and $\hat{x}_k \triangleq \hat{x}_{k|k} = E\{x_k | Y_k\}$ is the estimate of x_k, satisfying the following Kalman filtering equations

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k (y_k - C\hat{x}_{k|k-1}) \tag{11}$$
$$\hat{x}_{k|k-1} = A\hat{x}_{k-1|k-1} + Bu_{k-1} \tag{12}$$
$$K_k = \overline{P}_{k|k-1} C^T (C\overline{P}_{k|k-1} C^T + V)^{-1} \tag{13}$$
$$\overline{P}_{k|k-1} = A\overline{P}_{k-1|k-1} A^T + DUD^T \tag{14}$$
$$\overline{P}_{k|k} = \overline{P}_{k|k-1} - \overline{P}_{k|k-1} C^T (C\overline{P}_{k|k-1} C^T + V)^{-1} C\overline{P}_{k|k-1} \tag{15}$$

where $\overline{P}_{k|k} = E\{(x_{k|k} - \hat{x}_{k|k})(x_{k|k} - \hat{x}_{k|k})^T\}, \overline{P}_{0|0} = \overline{P}_0, \hat{x}_0 = m_0, Y_k = \{y_0, y_1, \cdots, y_k\}$.

Proof. The proof of Lemma 1 is similar to the LQG control of discrete stochastic systems in [16] and thus is omitted for brevity. □

In this paper, our objective is to develop a novel WPD-based real-time transcale LQG tracking control algorithm, which can ensure the multiscale processing algorithm be better combined with LQG tracking control algorithm than the WT-based algorithm in [45].

3. The WPD of discrete sequences

In the WT of discrete sequences, the approximations are the low-frequency components of discrete sequences and the details are the high-frequency components. For a given discrete sequence $a_{j,n} \in l_2$ at scale j ($a_{j,n}$ is assumed to be a scalar sequence here), the approximation coefficients at scale $j+1$ are derived by down sampling the output of low-pass filter by two

$$a_{j+1,k} = \sum_n h_{n-2k} a_{j,n} \tag{16}$$

Some details are lost from $a_{j,n}$ due to low-pass filtering, which could be computed by down sampling the output of high-pass filter by two

$$\delta a_{j+1,k} = \sum_n g_{n-2k} a_{j,n} \tag{17}$$

where $\delta a_{j+1,k}$ is called the detail coefficient at scale $j+1$. Here, h_k and g_k are the wavelet filter coefficients, which are determined by the scale function and wavelet function. Eqs. (16) and (17) are the decompositions of original discrete sequence, and the reconstruction is given as

$$a_{j,n} = \sum_k h_{n-2k} a_{j+1,k} + \sum_k g_{n-2k} \delta a_{j+1,k} \tag{18}$$

The WPD of discrete sequences is a transform where the discrete sequences are passed through more filters than WT. In the WT, each level is calculated by passing the previous approximation coefficients through high and low-pass filters. However, in the WPD, both the detail and approximation coefficients are decomposed [35]. The deduction equations of WPC are defined as

$$a_{j+1,k}^{2l} = \sum_n h_{n-2k} a_{j,n}^{l} \tag{19}$$

$$a_{j+1,k}^{2l+1} = \sum_n g_{n-2k} a_{j,n}^l \tag{20}$$

and the reconstruction is given as

$$a_{j,n}^l = \sum_k h_{n-2k} a_{j+1,k}^{2l} + \sum_k g_{n-2k} a_{j+1,k}^{2l+1} \tag{21}$$

where l represents the space serial number at scale j. Fig. 1 illustrates the decomposition mechanism used in WPD.

For simplicity and efficiency, the Haar WPD will be used through out this paper. As in (19) and (20), the Haar WPD is given as

$$a_{j+1,k}^{2l} = \frac{\sqrt{2}}{2}\left(a_{j,2k}^l + a_{j,2k+1}^l\right) \tag{22}$$

$$a_{j+1,k}^{2l+1} = \frac{\sqrt{2}}{2}\left(a_{j,2k}^l - a_{j,2k+1}^l\right) \tag{23}$$

Denote $j = 0, 1, \ldots, J$, where 0 and J represent the finest and the coarsest scale, respectively. For J scales decomposition, the WPD produces 2^J sets of coefficients as opposed to $J+1$ sets for the WT. However, due to the down sampling process, the overall number of coefficients is still the same and there is no redundancy. In fact, the initial frequency band of the discrete sequence is equally divided into 2^J frequency subbands under J scales WPD.

Denote k_j as the sequential time index with the sampling rate s_j at scale j, where s_j *satisfies* $s_j = \frac{s_0}{2^j}$ and k_j satisfies $k_j = 2^{J-j}k_J + 2^{J-j} - 1$, then the decomposition in (22) and (23) can be rewritten as

$$a_{j+1,k_{j+1}}^{2l} = \frac{\sqrt{2}}{2}\left(a_{j,k_j-1}^l + a_{j,k_j}^l\right) \tag{24}$$

$$a_{j+1,k_{j+1}}^{2l+1} = \frac{\sqrt{2}}{2}\left(a_{j,k_j-1}^l - a_{j,k_j}^l\right) \tag{25}$$

For a given discrete sequence with length $M2^J$ at scale 0, the J scale Haar WPD can be fulfilled in M data blocks, respectively. Fig. 2 illustrates the dyadic tree structure of WPC under 2 scale Haar WPD in the k_2th data block. Denote the following data block

$$\mathcal{A}^l(k_j) \triangleq \left[a_{j,k_j-2^{J-j}+1}^l, a_{j,k_j-2^{J-j}+2}^l, \cdots, a_{j,k_j}^l\right] \tag{26}$$

then the vector form of (24) and (25) can be derived in the following operator form

$$\begin{cases} \mathcal{A}^{2l}(k_{j+1}) = H^{j+1}\mathcal{A}^l(k_j) \\ A^{2l+1}(k_{j+1}) = G^{j+1}\mathcal{A}^l(k_j) \end{cases} \tag{27}$$

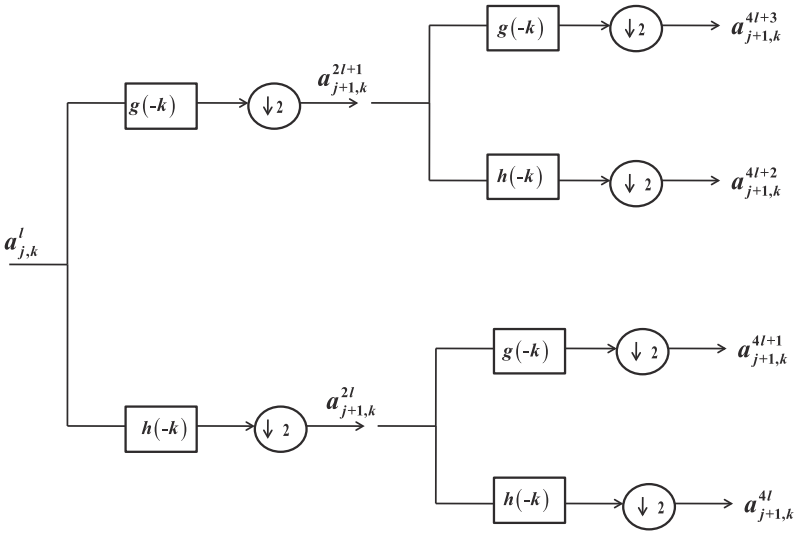

Fig. 1. Schematic diagrams of WPD.

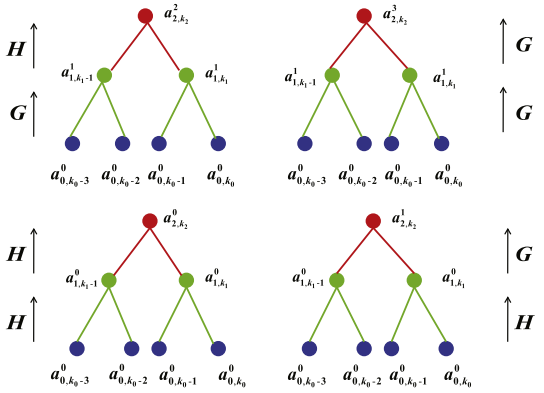

Fig. 2. The dyadic tree structure of WPC under 2 scale Haar WPD in the k_2th data block.

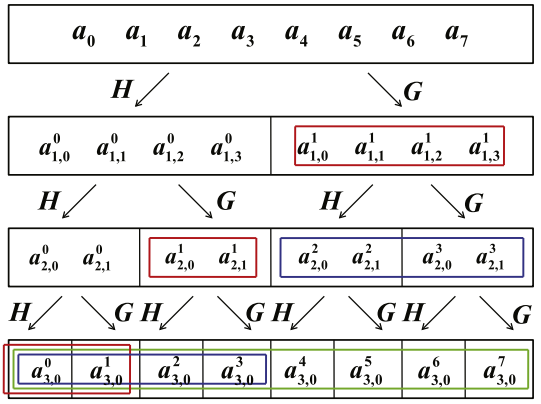

Fig. 3. The 3 scales WPD under different wavelet packet bases.

where operators $H^{j+1} \in \mathrm{R}^{2^{J-j-1}\times 2^{J-j}}$ *and* $G^{j+1} \in \mathrm{R}^{2^{J-j-1}\times 2^{J-j}}$ are composed of low-pass and high-pass Haar WPD filters coefficients from scale j to scale $j+1$. Moreover, $A^l(k_j)$ can be decomposed into the coarsest scale decomposition layer by using the following transformation

$$\begin{bmatrix} \mathcal{A}^{2^{J-j}l}(k_J) \\ \mathcal{A}^{2^{J-j}l+1}(k_J) \\ \vdots \\ \mathcal{A}^{2^{J-j}l+2^{J-j}-1}(k_J) \end{bmatrix} = T^{J|j}\mathcal{A}^l(k_j) \tag{28}$$

where $T^{J|j} = \begin{bmatrix} H^J H^{J-1}\cdots H^{j+1} \\ G^J H^{J-1}\cdots H^{j+1} \\ \vdots \\ G^J G^{J-1}\cdots G^{j+1} \end{bmatrix}$ is a square orthogonal matrix. For example, if $J=2$, we have

$$T^{2|0} = \begin{bmatrix} H^2H^1 \\ G^2H^1 \\ G^2G^1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \tag{29}$$

Remark 1. If $a^0_{0,k_0} \in \mathrm{R}^n$ is a vector sequence, the decomposition under Haar WPD still satisfies (24) and (25). For example, if $J=2$, we have

$$T^{2|0}=\begin{bmatrix}H^2H^1\\G^2H^1\\G^2G^1\end{bmatrix}=\begin{bmatrix}\frac{1}{2}I_2 & \frac{1}{2}I_2 & \frac{1}{2}I_2 & \frac{1}{2}I_2\\ \frac{1}{2}I_2 & \frac{1}{2}I_2 & -\frac{1}{2}I_2 & -\frac{1}{2}I_2\\ \frac{1}{2}I_2 & -\frac{1}{2}I_2 & \frac{1}{2}I_2 & -\frac{1}{2}I_2\\ \frac{1}{2}I_2 & -\frac{1}{2}I_2 & -\frac{1}{2}I_2 & \frac{1}{2}I_2\end{bmatrix} \tag{30}$$

where $I_2\in \mathrm{R}^{2\times 2}$ is a unit matrix.

Consider a discrete sequence $a=\{a_0,a_1,a_2,a_3,a_4,a_5,a_6,a_7\}$ where $a_i(i=0,1,\ldots,7)$ is constant. Denote Φ be an orthogonal basis and Φa be the sequence of coefficients of a under the decomposition of orthogonal basis Φ, then 3 scales WPD of a under different wavelet packet bases is shown in Fig. 3, where H and G are composed by filter response of Haar wavelet [13]. Here, $\Phi_1 a=\{a^0_{3,0},a^1_{3,0},a^2_{3,0},a^3_{3,0},a^4_{3,0},a^5_{3,0},a^6_{3,0},a^7_{3,0}\}$ is the coefficients of a in the wavelet packet basis, $\Phi_2 a=\{a^0_{3,0},a^1_{3,0},a^2_{3,0},a^3_{3,0},a^2_{2,0},a^2_{2,1},a^3_{2,0},a^3_{2,1}\}$ is the coefficients of a in another wavelet packet basis, $\Phi_3 a=\{a^0_{3,0},a^1_{3,0},a^1_{2,0},a^1_{2,1},a^1_{1,0},a^1_{1,1},a^1_{1,2},a^1_{1,3}\}$ is the coefficients of a in the wavelet basis, and Φ_1,Φ_2,Φ_3 are all orthogonal bases. In fact, we have other combination forms of coefficients under different wavelet packet bases, and the wavelet basis decomposition based WT is only a special WPD [13].

The best-basis selection is important in WPD. Ref. [13] applied an additive information cost function to ensure the best-basis such that information cost function is minimal. For different purposes, we can choose different orthogonal bases to act as best-basis and fulfill decomposition. In this paper, we choose the decomposition structure in $\Phi_1 a$ (all the coefficients at the coarsest scale decomposition layer) as the best basis decomposition and develop the algorithm in that case, since from the point of view of computing efficiency, this decomposition structure is better than others, which will be discussed later.

Remark 2. In this subsection, we have not given the further explanations about the decomposition and reconstruct equations in (16)–(23) for brevity. Readers can refer to [23,35,3,24] to get the detailed derivation of these equations.

4. Main results

Represent the system (1)–(4) into the following multiscale manner

$$x^0_{0,k_0+1}=Ax^0_{0,k_0}+Bu^0_{0,k_0}+Dw^0_{0,k_0} \tag{31}$$
$$y^0_{0,k_0}=Cx^0_{0,k_0}+v^0_{0,k_0} \tag{32}$$
$$z^0_{0,k_0}=C_d x^0_{0,k_0} \tag{33}$$
$$e^0_{0,k_0}=\bar{z}^0_{0,k_0}-z^0_{0,k_0} \tag{34}$$

Denote $A^{(j)}=A^{2^j}, B^{(j)}=(I+A^{2^{j-1}})\cdots(I+A)=\sum_{i=0}^{2^j-1}A^i, A^{(0)}=A$, and $B^{(0)}=I$, where $j=1,2,\ldots,J$. In order to facilitate the control development, the following assumptions are made for system (31)–(34):

Assumption 1. Matrix $B^{(j)}$ is nonsingular, $\mathrm{rank}(B)=m$, and $\mathrm{rank}(D)=p$.

Assumption 2. The matrix pair $(A^{(j)},C_d)$ is observable.

Remark 3. If all the eigenvalues of A being within the unit circle, then, for any j, $B^{(j)}$ is nonsingular.

4.1. The state-space models of WPC

The state-space models of WPC at the decomposition layer of scale j are established in this subsection.

Theorem 1. *Based on the Haar WPD, the state-space models of WPC at the decomposition layer of scale j are established as*

$$x^{l(j)}_{j,k_j+1}=A^{(j)}x^{l(j)}_{j,k_j}+B^{(j)}Bu^{l(j)}_{j,k_j}+B^{(j)}Dw^{l(j)}_{j,k_j} \tag{35}$$
$$y^{l(j)}_{j,k_j}=Cx^{l(j)}_{j,k_j}+v^{l(j)}_{j,k_j} \tag{36}$$
$$z^{l(j)}_{j,k_j}=C_d x^{l(j)}_{j,k_j} \tag{37}$$
$$x^{l(j)+1}_{j,k_j+1}=A^{(j)}x^{l(j)+1}_{j,k_j}+B^{(j)}Bu^{l(j)+1}_{j,k_j}+B^{(j)}Dw^{l(j)+1}_{j,k_j} \tag{38}$$
$$y^{l(j)+1}_{j,k_j}=Cx^{l(j)+1}_{j,k_j}+v^{l(j)+1}_{j,k_j} \tag{39}$$
$$z^{l(j)+1}_{j,k_j}=C_d x^{l(j)+1}_{j,k_j} \tag{40}$$

where $l(j)=0,2,\ldots,2^j-2$ is even sequence and

$$Bu_{j,k_j}^{l(j)} = \frac{(B^{(j)})^{-1}}{\sqrt{2}}\left[A^{(j-1)}B^{(j-1)}Bu_{j-1,k_{j-1}-1}^{\frac{l(j)}{2}} + (I+A^{(j-1)})B^{(j-1)}Bu_{j-1,k_{j-1}}^{\frac{l(j)}{2}} + B^{(j-1)}Bu_{j-1,k_{j-1}+1}^{\frac{l(j)}{2}}\right] \tag{41}$$

$$Dw_{j,k_j}^{l(j)} = \frac{(B^{(j)})^{-1}}{\sqrt{2}}\left[A^{(j-1)}B^{(j-1)}Dw_{j-1,k_{j-1}-1}^{\frac{l(j)}{2}} + (I+A^{(j-1)})B^{(j-1)}Dw_{j-1,k_{j-1}}^{\frac{l(j)}{2}} + B^{(j-1)}Dw_{j-1,k_{j-1}+1}^{\frac{l(j)}{2}}\right] \tag{42}$$

$$Bu_{j,k_j}^{l(j)+1} = \frac{(B^{(j)})^{-1}}{\sqrt{2}}\left[A^{(j-1)}B^{(j-1)}Bu_{j-1,k_{j-1}-1}^{\frac{l(j)}{2}} + (I-A^{(j-1)})B^{(j-1)}Bu_{j-1,k_{j-1}}^{\frac{l(j)}{2}} - B^{(j-1)}Bu_{j-1,k_{j-1}+1}^{\frac{l(j)}{2}}\right] \tag{43}$$

$$Dw_{j,k_j}^{l(j)+1} = \frac{(B^{(j)})^{-1}}{\sqrt{2}}\left[A^{(j-1)}B^{(j-1)}Dw_{j-1,k_{j-1}-1}^{\frac{l(j)}{2}} + (I-A^{(j-1)})B^{(j-1)}Dw_{j-1,k_{j-1}}^{\frac{l(j)}{2}} - B^{(j-1)}Dw_{j-1,k_{j-1}+1}^{\frac{l(j)}{2}}\right] \tag{44}$$

$$E\{w_{j,k_j}^{l(j)}\} = 0, E\{w_{j,k_j}^{l(j)}w_{j,k_j}^{l(j),T}\} = U_j^{l(j)}, E\{v_{j,k_j}^{l(j)}\} = 0, E\{v_{j,k_j}^{l(j)}v_{j,k_j}^{l(j),T}\} = V$$

$$E\{w_{j,k_j}^{l(j)}v_{j,k_j}^{l(j),T}\} = 0, U_j^{l(j)} = \frac{1}{2}\Big\{\left[D^+(B^{(j)})^{-1}A^{(j-1)}B^{(j-1)}D\right]U_{j-1}^{\frac{l(j)}{2}}\left[D^+(B^{(j)})^{-1}A^{(j-1)}B^{(j-1)}D\right]^T$$
$$+\left[D^+(B^{(j)})^{-1}(I+A^{(j-1)})B^{(j-1)}D\right]U_{j-1}^{\frac{l(j)}{2}}\left[D^+(B^{(j)})^{-1}(I+A^{(j-1)})B^{(j-1)}D\right]^T$$
$$+\left[D^+(B^{(j)})^{-1}B^{(j-1)}D\right]U_{j-1}^{\frac{l(j)}{2}}\left[D^+(B^{(j)})^{-1}B^{(j-1)}D\right]^T\Big\},$$

$$E\{w_{j,k_j}^{l(j)+1}\} = 0, E\{w_{j,k_j}^{l(j)+1}w_{j,k_j}^{l(j)+1,T}\} = U_j^{l(j)+1}$$

$$E\{v_{j,k_j}^{l(j)+1}\} = 0, E\{v_{j,k_j}^{l(j)+1}v_{j,k_j}^{T}\} = V, E\{w_{j,k_j}v_{j,k_j}^{l(j)+1,T}\} = 0$$

$$U_j^{l(j)+1} = \frac{1}{2}\Big\{\left[D^+(B^j)^{-1}A^{(j-1)}B^{(j-1)}D\right]U_{j-1}^{\frac{l(j)}{2}}\left[D^+(B^{(j)})^{-1}A^{(j-1)}B^{(j-1)}D\right]^T$$
$$+\left[D^+(B^j)^{-1}(I-A^{(j-1)})B^{(j-1)}D\right]U_{j-1}^{\frac{l(j)}{2}}\left[D^+(B^j)^{-1}(I-A^{(j-1)})B^{(j-1)}D\right]^T$$
$$+\left[D^+(B^j)^{-1}B^{(j-1)}D\right]U_{j-1}^{\frac{l(j)}{2}}\left[D^+(B^j)^{-1}B^{(j-1)}D\right]^T\Big\}, D^+ = (D^TD)^{-1}D^T.$$

Here, if $\frac{l(j)}{2}$ is even, $\frac{l(j)}{2} = l(j-1)$, otherwise if $\frac{l(j)}{2}$ is odd, $\frac{l(j)}{2} = l(j-1)+1, l(-1) = 0, U_0^0 = U$.

Proof. See the Appendix A. □

Note that the time models (31)–(33) at the finest scale 0 are transformed into the time-scale models (35)–(44) at different coarser scale decomposition layers. The outputs $z_{j,k_j}^{l(j)}$ and $z_{j,k_j}^{l(j)+1}$ are the WPC of output (3). Then, decompose the desired output as in (24) and (25), we obtain $\bar{z}_{j,k_j}^{l(j)}$ and $\bar{z}_{j,k_j}^{l(j)+1}$. Denote

$$\begin{cases} e_{j,k_j}^{l(j)} \triangleq z_{j,k_j}^{l(j)} - \bar{z}_{j,k_j}^{l(j)} \\ e_{j,k_j}^{l(j)+1} \triangleq z_{j,k_j}^{l(j)+1} - \bar{z}_{j,k_j}^{l(j)+1} \end{cases} \tag{45}$$

Lemma 2. *Based on the Haar WPD, the following equation can be derived,*

$$\sum_{k_0=0}^{N}\|e_{0,k_0}^0\|^2 = \sum_{l(J)}^{l(J)+1}\sum_{s=l(J)}^{M-1}\sum_{k_j=0}\|e_{j,k_j}^s\|^2 \tag{46}$$

where $N = M2^J - 1$ and $\|\cdot\|$ represents the 2-norm of vector.

Proof. See the Appendix B. □

Remark 4. Note that the tracking from z_{0,k_0}^0 to the desired output $\bar{z}_{0,k_0}^0$ at scale 0 is divided into 2^J independent tracking at the decomposition layer of scale J, which means if $z_{J,k_j}^{l(J)}$ and $z_{J,k_j}^{l(J)+1}$ can all "near" $\bar{z}_{J,k_j}^{l(J)}$ and $\bar{z}_{J,k_j}^{l(J)+1}$ respectively, the output z_{0,k_0}^0 can also "near" the desired output $\bar{z}_{0,k_0}^0$.

Remark 5. The multiscale system in (35)–(40) is a rather complex system involving three indexes: including a space serial number index, a scale index and a time index. For example, consider $x_{j,k_j}^{l(j)}$ in (35), $l(j)$ represents the space serial number index, j represents the scale index and k_j represents the time index.

4.2. The WPD-based algorithm

Based on Theorem 1, the 2^J state-space models for WPC at different sets can be established. The WPD-based real-time transcale tracking control algorithm relies on the following two parts.

Part 1: Design $u_{J,k_J}^{l(J)}$ *and* $u_{J,k_J}^{l(J)+1}$ at the decomposition layer of scale J.

Based on Lemma 1 and Theorem 1, designing $u_{J,k_J}^{l(J)}$ and $u_{J,k_J}^{l(J)+1}$ at the decomposition layer of scale J such that the following cost functions are minimized

$$J^{l(J)}(u_{J,k_J}^{l(J)}) \triangleq \frac{1}{2}E\left\{e_{J,M-1}^{l(J),T}\widehat{F}_J^{l(J)}e_{J,M-1}^{l(J)} + \sum_{k_J=0}^{M-2}\left(e_{J,k_J}^{l(J),T}\widehat{Q}_J^{l(J)}e_{J,k_J}^{l(J)} + u_{J,k_J}^{l(J),T}\widehat{R}_J^{l(J)}u_{J,k_J}^{l(J)}\right)\right\} \tag{47}$$

$$J^{l(J)+1}(u_{J,k_J}^{l(J)+1}) \triangleq \frac{1}{2}E\left\{e_{J,M-1}^{l(J)+1,T}\widehat{F}_J^{l(J)+1}e_{J,M-1}^{l(J)+1} + \sum_{k_J=0}^{M-2}\left(e_{J,k_J}^{l(J)+1,T}\widehat{Q}_J^{l(J)+1}e_{J,k_J}^{l(J)+1} + u_{J,k_J}^{l(J)+1,T}\widehat{R}_J^{l(J)+1}u_{J,k_J}^{l(J)+1}\right)\right\} \tag{48}$$

$\hat{F}_J^{l(J)}, \hat{Q}_J^{l(J)}, \hat{F}_J^{l(J)+1}, \hat{Q}_J^{l(J)+1} \in \mathrm{R}^{s\times s}, \hat{R}_J^{l(J)}, \hat{R}_J^{l(J)+1} \in \mathrm{R}^{m\times m}$ are positive definite matrices.

Remark 6. The weighting matrices at different sets are mutual independence, which are generally taken as diagonal matrices in engineering practice [1].

Part 2: Compute u_{0,k_0}^0 by using $u_{J,k_J}^{l(J)}$ and $u_{J,k_J}^{l(J)+1}$.

Set $u_{0,0}^0 = \beta_0^0, u_{0,0}^{l(1)} = \beta_0^{l(1)}, u_{0,0}^{l(1)+1} = \beta_0^{l(1)+1}, \ldots, u_{J-1,0}^{l(J-1)} = \beta_{J-1}^{l(J-1)}, u_{J-1,0}^{l(J-1)+1} = \beta_{J-1}^{l(J-1)+1}$, where $\beta_0^0 \in \mathrm{R}^m, \beta_0^{l(1)} \in \mathrm{R}^m, \beta_0^{l(1)+1} \in \mathrm{R}^m, \ldots, \beta_{J-1}^{l(J-1)} \in \mathrm{R}^m \beta_{J-1}^{l(J-1)+1} \in \mathrm{R}^m$ are given real-vectors. From (41) and (43), we have

$$\sqrt{2}B^{(j)}Bu_{j,k_j}^{l(j)} = A^{(j-1)}B^{(j-1)}Bu_{j-1,k_{j-1}-1}^{\frac{l(j)}{2}} + (I + A^{(j-1)})B^{(j-1)}Bu_{j-1,k_{j-1}}^{\frac{l(j)}{2}} + B^{(j-1)}Bu_{j-1,k_{j-1}+1}^{\frac{l(j)}{2}} \tag{49}$$

$$\sqrt{2}B^{(j)}Bu_{j,k_j}^{l(j)+1} = A^{(j-1)}B^{(j-1)}Bu_{j-1,k_{j-1}-1}^{\frac{l(j)}{2}} + (I - A^{(j-1)})B^{(j-1)}Bu_{j-1,k_{j-1}}^{\frac{l(j)}{2}} - B^{(j-1)}Bu_{j-1,k_{j-1}+1}^{\frac{l(j)}{2}} \tag{50}$$

Then, from (49) plus (50), we can obtain

$$u_{j-1,k_{j-1}}^{\frac{l(j)}{2}} = \frac{\sqrt{2}}{2}B^+(B^{(j-1)})^{-1}B^{(j)}B(u_{j,k_j}^{l(j)} + u_{j,k_j}^{l(j)+1}) - B^+(B^{(j-1)})^{-1}A^{(j-1)}B^{(j-1)}Bu_{j-1,k_{j-1}-1}^{\frac{l(j)}{2}} \tag{51}$$

from (49) minus (50), we can obtain

$$u_{j-1,k_{j-1}+1}^{\frac{l(j)}{2}} = \frac{\sqrt{2}}{2}B^+(B^{(j-1)})^{-1}B^{(j)}B(u_{j,k_j}^{l(j)} - u_{j,k_j}^{l(j)+1}) - B^+(B^{(j-1)})^{-1}A^{(j-1)}B^{(j-1)}Bu_{j-1,k_{j-1}}^{\frac{l(j)}{2}} \tag{52}$$

where $B^+ = (B^TB)^{-1}B^T$. Since $u_{j,0}^{l(j)}, u_{j,0}^{l(j)+1}$ are known, $u_{j-1,0}^{l(j-1)} = \beta_{j-1}^{l(j-1)}$, and $u_{j-1,0}^{l(j-1)+1} = \beta_{j-1}^{l(j-1)+1}, u_{j-1,1}^{\frac{l(j)}{2}}, u_{j-1,2}^{\frac{l(j)}{2}}$ can be computed from (51) and (52), $u_{j-1,3}^{\frac{l(j)}{2}}, u_{j-1,4}^{\frac{l(j)}{2}}$ can be computed from $u_{j,1}^{l(j)}, u_{j,1}^{l(j)+1}$ and $u_{j-1,2}^{\frac{l(j)}{2}}$, and so forth, $u_{j-1,k_{j-1}}^{\frac{l(j)}{2}}(k_{j-1} = 5, \ldots, 2M-2)$ can be successively computed, where if $\frac{l(j)}{2}$ is even, $\frac{l(j)}{2} = l(J-1)$, otherwise if $\frac{l(j)}{2}$ is odd, $\frac{l(j)}{2} = l(J-1)+1$. Then, repeat this process, u_{0,k_0}^0 can be computed from $u_{1,k_1}^{l(1)}, u_{1,k_1}^{l(1)+1}$.

The term real-time means that a control law is given at the finest scale at the previous sampling time, a single new measurement is acquired at the present sampling time, then the new state estimate information at the coarsest scale decomposition layer can be derived based upon all measurements available at the finest scale and the new control law is obtained at the finest scale at the present sampling time. Based on Parts 1–2, the WPD-based transcale LQG tracking control algorithm is developed as follows:

Step 1: Establish the 2^J state-space models of WPC at the coarsest scale decomposition layer J as in Theorem 1 and decompose the desired output.

Step 2: Solve the 2^J Riccati-type matrix difference equations as in (7) and (8), 2^J linear vector difference equations as in (9) and (10), and 2^J Kalman filtering gain equations as in (13)–(15) for different WPC sets at the coarsest scale decomposition layer J.

At the coarsest scale decomposition layer J, for $l(J)$, solve the Riccati-type matrix difference equations, linear vector difference equations, and the Kalman filtering gain equations as follows:

$$P_{J,k_J}^{l(J)} = C_d^T \widehat{Q}_J^{l(J)} C_d + A^{(J),T} P_{J,k_J+1}^{l(J)} A^{(J)} - A^{(J),T} P_{J,k_J+1}^{l(J)} B^{(J)} B[\widehat{R}_J^{l(J)} + (B^{(J)}B)^T P_{J,k_J+1}^{l(J)} B^{(J)} B]^{-1} (B^{(J)}B)^T P_{J,k_J+1}^{l(J)} A^{(J)} \tag{53}$$

$$P_{J,M-1}^{l(J)} = \widehat{F}_J^{l(J)} \tag{54}$$

$$g_{J,k_J}^{l(J)} = -C_d^T \widehat{Q}_J^{l(J)} \bar{z}_{J,k_J}^{l(J)} - A^{(J),T} P_{J,k_J+1}^{l(J)} B^{(J)} B[\widehat{R}_J^{l(J)} + (B^{(J)}B)^T P_{J,k_J+1}^{l(J)} B^{(J)} B]^{-1} (B^{(J)}B)^T g_{J,k_J+1}^{l(J)} + A^{(J),T} g_{J,k_J+1}^{l(J)} \tag{55}$$

$$g_{J,M-1}^{l(J)} = -C_d^T \widehat{F}_J^{l(J)} \bar{z}_{J,M-1}^{l(J)} \tag{56}$$

$$\overline{P}_{J,k_J|k_J-1}^{l(J)} = A^{(J)} \overline{P}_{J,k_J-1|k_J-1}^{l(J)} A^{(J),T} + D U_J^{l(J)} D^T \tag{57}$$

$$K_{J,k_J}^{l(J)} = \overline{P}_{J,k_J|k_J-1}^{l(J)} C^T [C \overline{P}_{J,k_J|k_J-1}^{l(J)} C^T + V]^{-1} \tag{58}$$

$$\overline{P}_{J,k_J|k_J}^{l(J)} = \overline{P}_{J,k_J|k_J-1}^{l(J)} - \overline{P}_{J,k_J|k_J-1}^{l(J)} C^T [C \overline{P}_{J,k_J|k_J-1}^{l(J)} C^T + V]^{-1} C \overline{P}_{J,k_J|k_J-1}^{l(J)} \tag{59}$$

for $l(J)+1$, solve the Riccati-type matrix difference equations, linear vector difference equations, and the Kalman filtering gain equations as follows:

$$P_{J,k_J}^{l(J)+1} = C_d^T \widehat{Q}_J^{l(J)+1} C_d + A^{(J),T} P_{J,k_J+1}^{l(J)+1} A^{(J)} - A^{(J),T} P_{J,k_J+1}^{l(J)+1} B^{(J)} B[\widehat{R}_J^{l(J)+1} + (B^{(J)}B)^T P_{J,k_J+1}^{l(J)+1} B^{(J)} B]^{-1} (B^{(J)}B)^T P_{J,k_J+1}^{l(J)+1} A^{(J)} \tag{60}$$

$$P_{J,M-1}^{l(J)+1} = \widehat{F}_J^{l(J)+1} \tag{61}$$

$$g_{J,k_J}^{l(J)+1} = -C_d^T \widehat{Q}_J^{l(J)+1} \bar{z}_{J,k_J}^{l(J)} - A^{(J),T} P_{J,k_J+1}^{l(J)} B^{(J)} B[\widehat{R}_J^{l(J)+1} + (B^{(J)}B)^T P_{J,k_J+1}^{l(J)+1} B^{(J)} B]^{-1} (B^{(J)}B)^T g_{J,k_J+1}^{l(J)+1} + A^{(J),T} g_{J,k_J+1}^{l(J)+1} \tag{62}$$

$$g_{J,M-1}^{l(J)+1} = -C_d^T \widehat{F}_J^{l(J)+1} \bar{z}_{J,M-1}^{l(J)+1} \tag{63}$$

$$\overline{P}_{J,k_J|k_J-1}^{l(J)+1} = A^{(J)} \overline{P}_{J,k_J-1|k_J-1}^{l(J)+1} A^{(J),T} + D U_J^{l(J)+1} D^T \tag{64}$$

$$K_{J,k_J}^{l(J)+1} = \overline{P}_{J,k_J|k_J-1}^{l(J)+1} C^T [C \overline{P}_{J,k_J|k_J-1}^{l(J)+1} C^T + V]^{-1} \tag{65}$$

$$\overline{P}_{J,k_J|k_J}^{l(J)+1} = \overline{P}_{J,k_J|k_J-1}^{l(J)+1} - \overline{P}_{J,k_J|k_J-1}^{l(J)+1} C^T [C \overline{P}_{J,k_J|k_J-1}^{l(J)+1} C^T + V]^{-1} C \overline{P}_{J,k_J|k_J-1}^{l(J)+1} \tag{66}$$

where $l(J) = 0, 2, \ldots, 2^J - 2$.

Step 3: Compute the control law u_{0,k_0}^0 at scale 0.

At the decomposition layer of scale J, assume that $\hat{x}_{J,k_J-1|k_J-1}^{l(J)}, \hat{x}_{J,k_J-1|k_J-1}^{l(J)+1}$ are given, then from Steps 1–2, we have

$$\hat{x}_{J,k_J|k_J-1}^{l(J)} = A^{(J)} \hat{x}_{J,k_J-1|k_J-1}^{l(J)} + B^{(J)} B u_{J,k_J-1}^{l(J)} \tag{67}$$

$$\hat{x}_{J,k_J|k_J}^{l(J)} = \hat{x}_{J,k_J|k_J-1}^{l(J)} + K_{J,k_J}^{l(J)} (y_{J,k_J}^{l(J)} - C \hat{x}_{J,k_J|k_J-1}^{l(J)}) \tag{68}$$

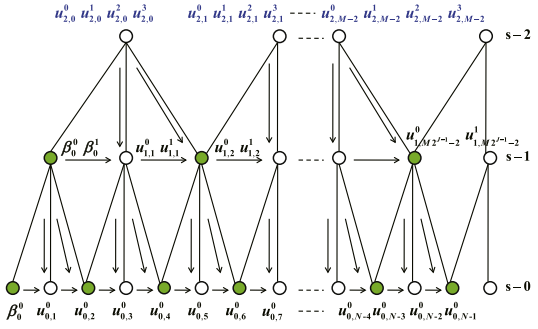

Fig. 4. Compute u_{0,k_0}^0 under 2 scales Haar WPD.

$$u_{J,k_J}^{l(J)} = -[\widehat{R}_J^{l(J)} + (B^{(J)}B)^T P_{J,k_J+1}^{l(J)} B^{(J)}B]^{-1} (B^{(J)}B)^T (P_{J,k_J+1}^{l(J)} A^{(J)} \hat{x}_{J,k_J|k_J}^{l(J)} + g_{J,k_J+1}^{l(J)}) \tag{69}$$

$$\hat{x}_{J,k_J|k_J-1}^{l(J)+1} = A^{(J)} \hat{x}_{J,k_J-1|k_J-1}^{l(J)+1} + B^{(J)} B u_{J,k_J-1}^{l(J)+1} \tag{70}$$

$$\hat{x}_{J,k_J|k_J}^{l(J)+1} = \hat{x}_{J,k_J|k_J-1}^{l(J)+1} + K_{J,k_J}^{l(J)+1} (y_{J,k_J}^{l(J)+1} - C\hat{x}_{J,k_J|k_J-1}^{l(J)+1}) \tag{71}$$

$$u_{J,k_J}^{l(J)+1} = -[\widehat{R}_J^{l(J)+1} + (B^{(J)}B)^T P_{J,k_J+1}^{l(J)+1} B^{(J)}B]^{-1} (B^{(J)}B)^T (P_{J,k_J+1}^{l(J)+1} A^{(J)} \times \hat{x}_{J,k_J|k_J}^{l(J)+1} + g_{J,k_J+1}^{l(J)}) \tag{72}$$

and we can use $u_{J,k_J}^{l(J)}$ and $u_{J,k_J}^{l(J)+1}$ to compute $u_{0,k_0}^0, u_{0,k_0+1}^0, \ldots, u_{0,k_0+2^J-1}^0$ as in Part 2, thus we have $y_{0,k_0+1}^0, y_{0,k_0+2}^0, \ldots, y_{0,k_0+2^J}^0$, which means $y_{J,k_J+1}^{l(J)}$ and $y_{J,k_J+1}^{l(J)+1}$ can be obtained, then repeat the above process. Fig. 4 shows one completely real-time computing process of the control law u_{0,k_0}^0 at scale 0 under $J = 2$.

4.2.1. Computational efficiency analysis

4.2.1.1. The LQG tracking control algorithm (LQGTC). In the LQG tracking control algorithm in Lemma 1, since the tracking time interval is defined in $[0, N]$, the computational numbers for Riccati-type matrix difference equations in (7) and (8), linear vector difference equations in (9) and (10) and Kalman filtering Eqs. (13)–(17) are all $O(N)$.

4.2.1.2. The WPD-based transcale LQG tracking control algorithm (WPD-TLQGTC). In the WPD-based transcale LQG tracking control algorithm, for J scales decomposition, the WPD produces 2^J sets of coefficients, and based on the features of Haar WPD [35], the 2^J sets of coefficients can be fulfilled in parallel, which means the computation for the Riccati-type matrix difference equations, linear vector difference equations, and Kalman filtering gain equations in each coefficients set in Step 2 can run in parallel, i.e., the computing of (53)–(59) and (60)–(66) can run in parallel. The computational numbers for the Riccati-type matrix difference equations, linear vector difference equations, and Kalman filtering gain equations in each coefficients sets are all $O(N/2^J)$, which is the $1/2^J$ of that of the conventional LQG tracking control algorithm. In Step 3, the computation of (67)–(69) and (70)–(72) can also run in parallel, and the computational number for each parallel part is $O(N/2^J)$.

4.2.1.3. The WT-based transcale LQG tracking control algorithm in [45] (WT-TLQGTC). In the WT-based transcale LQG tracking control algorithm, for J scales decomposition, the WT produces $J+1$ sets of coefficients, and the $J+1$ Riccati-type matrix difference equations, $J+1$ linear vector difference equations, and $J+1$ Kalman filtering equations in different coefficients sets can run in parallel, and the maximum computational number is $O(N/2)$.

4.2.1.4. Comparison. Table 1 shows the comparison of the maximum computational number for the Riccati-type matrix difference equations (RTDEs), linear vector difference equations (LVDEs), and Kalman filtering equations (KFEs) in each parallel computation for the three control algorithms. For the conventional LQG tracking algorithm, it cannot be fulfilled in parallel, so the maximum computational number is $O(N)$. For the WPD-based transcale LQG tracking control algorithm, the 2^J sets of coefficients can be fulfilled in parallel, and the computational number for each set is the same, so the maximum computational number is $O(N/2^J)$. It is clearly shown from Table 1 that the maximum computational number for the WPD-based transcale LQG tracking control algorithm is the $1/2^J$ of that of the conventional LQG tracking control algorithm, and the $1/2^{J-1}$ of that of the WT-based transcale LQG tracking control algorithm in [45].

It should be pointed out that although the proposed WPD-based transcale LQG tracking control algorithm increases the computations from control laws $u_{J,k_J}^{l(J)}, u_{J,k_J}^{l(J)}$ to control law u_{0,k_0}^0 as in Part 2 compared with the LQG tracking control algorithm, since the computations in Part 2 are all additive and subtraction operations, these computations will not cost a lot of time. Based on the above analysis, the WPD-based transcale control algorithm further improves the computational efficiency compared with the conventional LQG tracking control algorithm than the WT based algorithm.

Remark 7. In this paper, we choose all the decomposition coefficients at the coarsest scale decomposition layer as the best basis decomposition and develop the algorithm in that case. If we choose other wavelet packet basis as the best basis, such as the wavelet basis, the processes in Steps 1–3 are similar, which means the WT-based transcale LQG tracking control

Table 1
Comparison of the three control algorithms.

Algorithm	RTDEs	LVDEs	KFEs
LQGTC	$O(N)$	$O(N)$	$O(N)$
WT-TLQGTC [45]	$O(N/2)$	$O(N/2)$	$O(N/2)$
WPD-TLQGTC	$O(N/2^J)$	$O(N/2^J)$	$O(N/2^J)$

algorithm in [45] is actually a special case of the WPD-based algorithm developed in this paper. From the point of the view of computation efficiency, the decomposition structure in this paper is better than that of others, because under this decomposition structure, the number of decomposition coefficients at each set is not only equal but also the fewest.

Remark 8. It can be seen from Steps 1–3 that although the computation number for each coefficients set can be further reduced if we choose larger J, the computing from control laws $u_{J,k_J}^{l(J)}, u_{J,k_J}^{l(J)}$ to control law u_{0,k_0}^0 will become more complex and more computing time will be cost, so how to decide the decomposition scale J in the algorithm should consider the compromise between the computing time of equations in (53)–(72) and the computing time of equations as in Part 2.

Remark 9. If the desired output is chosen as $\bar{z}_k = 0 \in \mathrm{R}^n$, the algorithm in Lemma 1 reduces to the conventional finite-horizon LQG control algorithm. Steps 1–3 can be applied to develop the WPD-based transcale LQG control algorithm. Moreover, if the system noise w_k and the measurement noise v_k are not considered, the algorithm in Lemma 1 reduces to the conventional finite-horizon LQ tracking control algorithm, Steps 1–3 can be applied to develop the WPD-based transcale LQ tracking control algorithm [44].

Denote

$$\begin{bmatrix} M_1^{(j)} & M_2^{(j)} \\ M_3^{(j)} & M_4^{(j)} \end{bmatrix} \triangleq M^{-1}\left(B^{(j-1)}\right)^{-1} B^{(j)} M^{-1,T} \tag{73}$$

where $M \in \mathrm{R}^{n\times n}$ is an elementary matrix such that $B^T M = [\,0 \quad I_m\,], I_m \in \mathrm{R}^{m\times m}$ is an unit matrix, $M_4^{(j)} \in \mathrm{R}^{m\times m}$.

Theorem 2. *If $M_4^{(j)}$ is nonsingular, the control law u_{0,k_0}^0 solved from Part 2 can guarantee the performance indexes (47) and (48) at the coarsest scale decomposition layer are minimized.*

Proof. The proof of Theorem 2 is similar to that of Theorem 2 in [45] and thus is omitted for brevity. □

Remark 10. From $\left(B^{(j-1)}\right)^{-1} B^{(j)} = I + A^{2^{j-1}}$, if all the eigenvalues of A being with in the unit circle, $M_4^{(j)}$ is nonsingular can be easily guaranteed.

Theorem 3. *If the system (1) and (2) satisfies Assumptions 1 and 2 and $M_4^{(j)}$ is nonsingular, the following inequality can be guaranteed under the proposed algorithm*

$$E\left\{\frac{1}{2}e_{0,N}^T F e_{0,N} + \frac{1}{2}\sum_{k_0=0}^{N-1} e_{0,k_0}^T Q e_{0,k_0}\right\} \leqslant \frac{\lambda_{\max}}{\lambda_{\min}} \sum_{l(J)} \sum_{s=l(J)}^{l(J)+1} J^{s,*}. \tag{74}$$

where

$$J^{l(J),*} \triangleq \frac{1}{2}E\left\{ e_{J,M-1}^{l(J),T} \widehat{F}_J^{l(J)} e_{J,M-1}^{l(J)} + \sum_{k_J=0}^{M-2} \left(e_{J,k_J}^{l(J),T} \widehat{Q}_J^{l(J)} e_{J,k_J}^{l(J)} \right) \right\} \tag{75}$$

$$J^{l(J)+1,*} \triangleq \frac{1}{2}E\left\{ e_{J,M-1}^{l(J)+1,T} \widehat{F}_J^{l(J)+1} e_{J,M-1}^{l(J)+1} + \sum_{k_J=0}^{M-2} (e_{J,k_J}^{l(J)+1,T} \widehat{Q}_J^{l(J)+1} e_{J,k_J}^{l(J)+1}) \right\} \tag{76}$$

represent the optimal performance indexes of tracking error part under the control laws $u_{J,k_J}^{l(J)}, u_{J,k_J}^{l(J)+1}$, and $\lambda_{\max} = \max\{\max\{\lambda(F)\}, \max\{\lambda(Q)\}\} \lambda_{\min} = \min\{\min\{\lambda(\widehat{F}_J^{l(J)})\}, \min\{\lambda(\widehat{Q}_J^{l(J)})\}, \min\{\lambda(\widehat{F}_J^{l(J)+1})\}, \min\{\lambda(\widehat{Q}_J^{l(J)+1})\}\}, \max\{\lambda(\bullet)\}$, $\min\{\lambda(\bullet)\}$ represent the maximum and minimum eigenvalues of matrix, respectively.

Proof. See the Appendix C. □

5. Illustrative example

Example 1. Consider the following linearized satellite attitude model with small angular maneuvers under the influence of gravity gradient torque around Local-Vertical–Local-Horizontal (LHLV) reference frame [20],

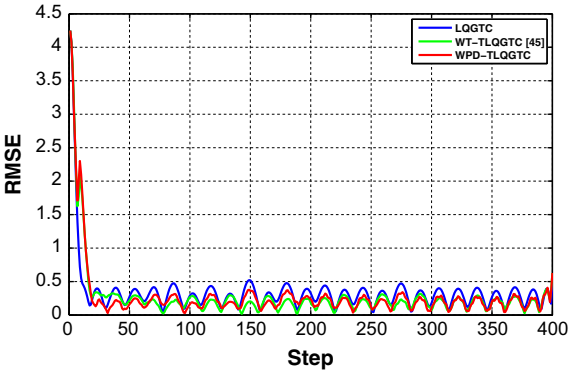

Fig. 5. RMSE in output.

$$\dot{x}_t = Gx_t + Hu_t + Nw_t \tag{77}$$
$$y_t = Cx_t + v_t \tag{78}$$

where $x_t = [\phi, \theta, \psi, \dot{\phi}, \dot{\theta}, \dot{\psi}] \in \mathrm{R}^6, u_t \in \mathrm{R}^3$ and $y_t \in \mathrm{R}^6$ are the state, input torque and measurement vectors respectively. G is the state matrix, H is control input matrix, N is disturbance input matrix and C is the measurement matrix. The system noise $w_t \in \mathrm{R}^3$ and the measurement noise $v_t \in \mathrm{R}^6$ are uncorrelated white zero-mean Gaussian vectors, where the statistical characters are given as

$$\begin{cases} E\{w_t\} = 0, E\{v_t\} = 0 \\ E\{w_t w_\tau^T\} = Q_e \delta_{t-\tau} \\ E\{v_t v_\tau^T\} = R_e \delta_{t-\tau} \\ E\{w_t v_\tau^T\} = 0 \end{cases} \tag{79}$$

In the particular problem they are given by

$$G = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \frac{4\varpi_0^2(J_z - J_y)}{J_x} & 0 & 0 & 0 & 0 & \frac{\varpi_0(J_x - J_y + J_z)}{J_x} \\ 0 & \frac{3\varpi_0^2(J_x - J_z)}{J_y} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\varpi_0^2(J_x - J_y)}{J_z} & \frac{\varpi_0(J_y - J_x - J_z)}{J_z} & 0 & 0 \end{bmatrix}$$

$$H = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \frac{1}{J_x} & 0 & 0 \\ 0 & \frac{1}{J_y} & 0 \\ 0 & 0 & \frac{1}{J_z} \end{bmatrix}, \quad C = I_{6\times 6}$$

where ϕ, θ, ψ are Eular angles, J_x, J_y, J_z are three axis inertia matrices, ϖ_0 is the mean orbital motion. The simulation parameters are chosen as in [20], where $J_x = 305.89126\ \mathrm{kg\,m^2}$, $J_y = 314.06488\ \mathrm{kg\,m^2}$, $J_z = 167.33919\ \mathrm{kg\,m^2}$, $\varpi_0 = 0.001$ rad/s, $\phi(0) = 3\mathrm{deg}$, $\theta = -1\mathrm{deg}$, $\psi = 0.5\mathrm{deg}$, $\dot{\phi}(0) = 0.1\mathrm{deg/s}$, $\dot{\theta} = -0.1\mathrm{deg/s}$, $\dot{\psi} = 0.1\mathrm{deg/s}$, $Q_e = \mathrm{diag}\{5\times10^{-3}, 5\times10^{-3}, 5\times10^{-3}\}$, $R_e = \mathrm{diag}\{10^{-2}, 10^{-2}, 10^{-2}\}$. It should be pointed out that the satellite attitude dynamic system given in (77) and (78) is a continuous model, however, the calculation of digital computer is discrete in time, when the system is controlled by digital computer or the controlled system is modeled, analyzed and designed by digital computer, the continuous model should be discretized. So the continuous model (77) and (78) should be discretized into discrete model as in (1)–(3), where $A = \exp(GT), B = (\int_0^T \exp(Gt)\mathrm{d}t)H, D = (\int_0^T \exp(Gt)\mathrm{d}t)N$, the sampling period T is chosen as 0.1 s, $z_k \in \mathrm{R}^3$ is defined as output vector and the output matrix is chosen as $C_d = I_3$. For convenience, the desired output is given as in discrete form

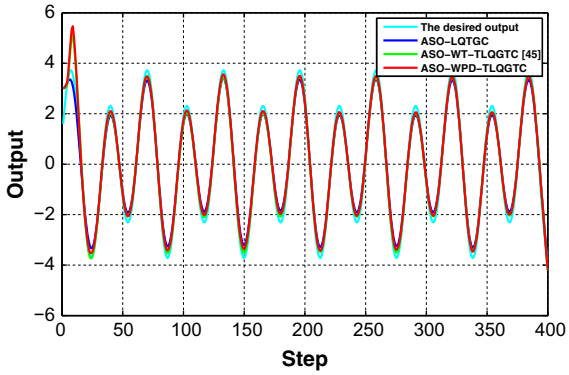

Fig. 6. The first component of desired output and ASO under LQGTC, WT-TLQGTC and WPD-TLQGTC respectively.

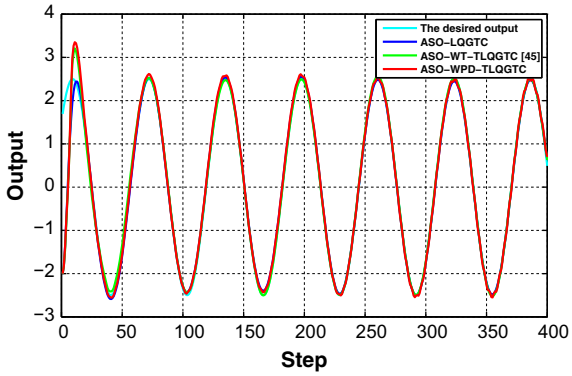

Fig. 7. The second component of desired output and ASO under LQGTC, WT-TLQGTC and WPD-TLQGTC respectively.

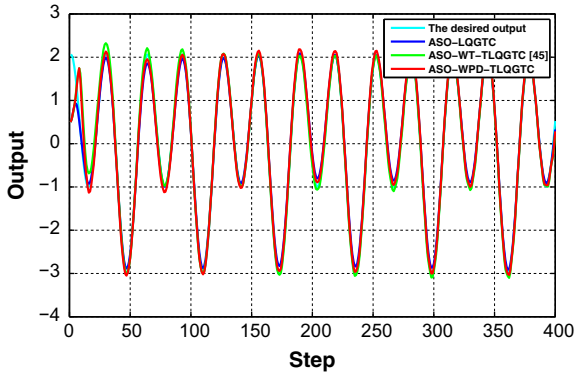

Fig. 8. The third component of desired output and ASO under LQGTC, WT-TLQGTC and WPD-TLQGTC respectively.

$$\bar{z}_k = \begin{bmatrix} 3\sin(0.2k) + \cos(0.1k) \\ 2\sin(0.1k) + 1.5\cos(0.1k) \\ 2\cos(0.2k) + \sin(0.1k) \end{bmatrix} \tag{80}$$

For comparison, the LQG tracking control algorithm in Lemma 1, WT-based transcale LQG tracking control algorithm in [45] and WPD-based transcale LQG tracking control algorithm in this paper will all be applied. For the sake of simplicity, we denote "LQGTC", "WT-TLQGTC" and "WPD-TLQGTC" for the LQG tracking control, WT-based transcale LQG tracking control and WPD-based transcale LQG tracking control, respectively. The coarsest scale is chosen as $J = 2$, the quadratic stochastic performance indexes in (47) and (48) are chosen at scale 2 and the weighting matrices in (47) and (48) are chosen as $\widehat{F}_2^0 = 1, \widehat{Q}_2^0 = 1, \widehat{R}_2^0 = 10^{-7}, \widehat{F}_2^2 = 1, \widehat{Q}_2^2 = 1, \widehat{R}_2^2 = 10^{-7}, \widehat{F}_2^1 = 1, \widehat{Q}_2^1 = 1, \widehat{R}_2^1 = 10^{-7}, \widehat{F}_2^3 = 1, \widehat{Q}_2^3 = 1, \widehat{R}_2^3 = 10^{-7}$, respectively. For

Table 2
Comparison of three control algorithms, where $J = 2, N = 400$.

Algorithm	API	CPU-time
LQGTC	375.1643	7.7385
WT-TLQGTC [45]	669.7347	4.0894
WPD-TLQGTC	713.0164	2.4585

Table 3
Comparison of three control algorithms, where $J = 3, N = 400$.

Algorithm	API	CPU-time
LQGTC	375.1643	7.7385
WT-TLQGTC [45]	926.2310	4.5178
WPD-TLQGTC	1023.2158	2.0476

WT-LQGTC in [45], the weighting matrices are chosen as $\widehat{F}_2^0 = 1, \widehat{Q}_2^0 = 1, \widetilde{R}_2^0 = 10^{-7}, \widetilde{F}_2^1 = 1, \widetilde{Q}_2^1 = 1, \widetilde{R}_2^1 = 10^{-7}, \widetilde{F}_1^1 = 1, \widetilde{Q}_1^1 = 1, \widetilde{R}_1^1 = 10^{-7}$, respectively, and for LQGTC, the weighting matrices are chosen as $F = 1, Q = 1$ and $R = 10^{-7}$. To compare the control performance, denote $z_{0,k_0}^{0,i}$ as the output at time k_0 at the ith Monte Carlo run. Suppose K of such Monte Carlo runs are carried out, the root mean square error (RMSE) in output at time k_0 can be calculated as [45]

$$\mathrm{RMSE}_{k_0} \triangleq \sqrt{\frac{1}{K}\sum_{i=1}^{K}\left[\left(z_{0,k_0}^{0,i} - \bar{z}_{0,k_0}\right)^T\left(z_{0,k_0}^{0,i} - \bar{z}_{0,k_0}\right)\right]} \tag{81}$$

The following average performance indexes (API) is also defined as performance metric, where

$$\mathrm{API} \triangleq \frac{1}{K}\sum_{i=1}^{K}\left[\frac{1}{2}(e_{0,N}^i)^T F e_{0,N}^i + \frac{1}{2}\sum_{k_0=0}^{N-1}(e_{0,k_0}^i)^T Q e_{0,k_0}^i + \frac{1}{2}\sum_{k_0=0}^{N-1}(u_{0,k_0}^i)^T R u_{0,k_0}^i\right] \tag{82}$$

and the subscript i in (81) and (82) represents the ith Monte Carlo run.

For $J = 2$, the simulation results averaged over 50 Monte Carlo runs are shown in Figs. 5–8 and Table 2. Fig. 5 shows the simulation results of the corresponding RMSE in output, which reflect the tracking errors. Figs. 6–8 show the simulation results of three components of desired output $\bar{z}_{0,k_0}$, average system output (ASO) $\frac{1}{K}\sum_{i=1}^{K} z_{0,k_0}^{0,i}$ under LQGTC, WT-TLQGTC and WPD-TLQGTC, respectively. The values of API and CPU-time of WPD-TLQGTC are compared in Table 2 with that of WT-LQGTC and LQGTC, respectively. It can be seen from Figs. 6–8 that the three algorithms can all provide good tracking performance. It is perhaps surprising to see from Fig. 5 that the tracking errors of WPD-TLQGTC and WT-LQGTC are better than that of LQGTC at some time, at least for this 2 scales decomposition case that we simulated. However, this dose not mean the WPD-TLQGTC algorithm or the WT-LQGTC algorithm with better performances compared with LQGTC, since the conventional LQGTC guarantees the integrate performance (including the energy of tracking errors part and controllers part) is optimal, but not guarantees each part is optimal. Fig. 5 only reflects the performance of tracking errors part in some extent. In fact, the performance of WPD-TLQGTC algorithm and WT-LQGTC algorithm are determined by choosing optimal indexes at different scales (by choosing F, Q, R at different scales), which gives us more chances to improve the performance of tracking errors part compared with conventional LQGTC. It is clearly seen from Table 2 that the integrate performances (the values of API) of WPD-TLQGTC and WT-LQGTC are worse than that of LQGTC since the LQGTC guarantees the performance optimal at the finest scale but the WPD-TLQGTC algorithm guarantees the performance optimal at the coarsest scale decomposition layer and the WT-TLQGTC algorithm in [45] guarantees the performance optimal at each coarser scale, respectively.

To assess the computation efficiency of the proposed method, we compute the total CPU time in MATLAB 7.10 on a 3.30 GHz i3-2120 CPU Intel Core-based computer operating under Windows XP (Professional). It appears that the average CPU time for WPD-TLQGTC is about 7.7385 s, that for WT-TLQGTC is about 4.0894 s and that for the LQGTC is about 2.4585 s. These accord to our analysis in computational efficiency part, that the CPU-time of the LQGTC algorithm is about one time longer than that of the WT-TLQGTC algorithm and the CPU-time of the WT-TLQGTC algorithm is about one time longer than that of the WPD-TLQGTC algorithm, which mean the WPD-TLQGTC algorithm proposed in this paper can provide better computation efficiency than that of WT-LQGTC algorithm in [45].

Moreover, for $J = 3$, the simulation results are shown in Table 3. Comparing with the CPU-time in Tables 2 and 3, which shows that, with the increasing of scale J, the time for WT-TLQGTC algorithm is increasing, and the time for WPD-TLQGTC algorithm is reducing, these according our analysis in computational efficiency part, since the computational efficiency for computing Riccati-type matrix difference equations, linear vector difference equations and Kalman filtering equations can be

further improved under WPD-TLQGTC algorithm with the increasing of scale J. It is seen from Tables 2 and 3 that the integrate performances (the values of API) of WPD-TLQGTC and WT-LQGTC under $J = 3$ are worse than that of under $J = 2$. These comparisons further show that the proposed WPD-TLQGTC algorithm can provide better computational performance than the WT-TLQGTC algorithm in [45].

Based on the above simulation results, we can conclude that the proposed WPD-TLQGTC algorithm is effective for discrete stochastic systems, further, it provides a better compromise between performance index and computational efficiency compared with the conventional LQG tracking control algorithm than the WT-TLQGTC algorithm in [45].

6. Conclusions

This paper introduces a Haar WPD-based real-time transcale LQG tracking control algorithm for the finite-time output tracking problem of discrete stochastic systems. Based on the state-space models of WPC established at the coarsest scale decomposition layer, the algorithm is developed by integrating two parts: one shows how to design the desired control laws at the coarsest scale decomposition layer, and the other shows how to recursively compute the control law at the finest scale. The simulation results show that the proposed algorithm not only compensates the phenomena that may occur at different scales but also provides a better compromise between performance index and computational efficiency compared with the conventional LQG tracking control algorithm than the WT-based algorithm in [45].

Acknowledgments

This work was supported by the National Basic Research Program of China (973 Program: 2012CB821200, 2012CB821201) and the National Nature Science Foundation of China (61134005, 61221061, 61327807, 61304232).

Appendix A. Proof of Theorem 1

To prove Theorem 1, the induction method is used. The state-space model at scale 0 is given as in (31)–(33). Use the Haar WPD, the state-space models of WPD at the decomposition layer of scale 1 can be generated as

$$\begin{aligned} x^0_{1,k_1+1} &= \frac{\sqrt{2}}{2}\left(x^0_{0,k_0+1} + x^0_{0,k_0+2}\right) \\ &= \frac{\sqrt{2}}{2}\left(A^2x^0_{0,k_0-1} + ABu^0_{0,k_0-1} + Bu^0_{0,k_0} + ADw^0_{0,k_0-1} + Dw^0_{0,k_0} + A^2x^0_{0,k_0} + ABu^0_{0,k_0} + Bu^0_{0,k_0+1} + ADw^0_{0,k_0} + Dw^0_{0,k_0+1}\right) \\ &= A^2x^0_{1,k_1} + (I+A)Bu^0_{1,k_1} + (I+A)Dw^0_{1,k_1} \end{aligned} \tag{83}$$

and

$$x^1_{1,k_1+1} = \frac{\sqrt{2}}{2}\left(x^0_{0,k_0+1} - x^0_{0,k_0+2}\right) = A^2x^1_{1,k_1} + (I+A)Bu^1_{1,k_1} + (I+A)Dw^1_{1,k_1} \tag{84}$$

At the decomposition layer of scale 1

$$y^1_{0,k_1} = \frac{\sqrt{2}}{2}\left(y^0_{0,k_0-1} + y^0_{0,k_0}\right) = \frac{\sqrt{2}}{2}\left(Cx^0_{0,k_0-1} + v^0_{0,k_0-1} + Cx^0_{0,k_0} + v^0_{0,k_0}\right) = Cx^1_{1,k_1} + v^1_{1,k_1} \tag{85}$$

$$y^1_{1,k_1} = \frac{\sqrt{2}}{2}\left(y^0_{0,k_0-1} - y^0_{0,k_0}\right) = Cx^1_{1,k_1} + v^1_{1,k_1} \tag{86}$$

$$z^0_{1,k_1} = \frac{\sqrt{2}}{2}\left(z^0_{0,k_0-1} + z^0_{0,k_0}\right) = \frac{\sqrt{2}}{2}\left(C_dx^0_{0,k_0-1} + C_dx^0_{0,k_0}\right) = C_dx^0_{1,k_1} \tag{87}$$

$$z^1_{1,k_1} = \frac{\sqrt{2}}{2}\left(z^0_{0,k_0-1} - z^0_{0,k_0}\right) = C_dx^1_{1,k_1} \tag{88}$$

thus the state-space models WPD at the decomposition layer of scale 1 are established. By using the linear property of expectation and Assumption 1, we can obtain $E\{w^0_{1,k_1}\} = 0, E\{w^0_{1,k_1}w^{0,T}_{1,k_1}\} = U^0_1, E\{v^0_{1,k_1}\} = 0, E\{v^0_{1,k_1}v^{0,T}_{1,k_1}\} = V, E\{w^0_{1,k_1}v^{0,T}_{1,k_1}\} = 0$ and $E\{w^1_{1,k_1}\} = 0, E\{w^1_{1,k_1}w^{1,T}_{1,k_1}\} = U^1_1, E\{v^1_{1,k_1}\} = 0, E\{v^1_{1,k_1}(v^1_{1,k_1})^T\} = V, E\{w^1_{1,k_1}v^{1,T}_{1,k_1}\} = 0$. Suppose that the state-space models at the decomposition layer of scale $j-1$ are of the form (35)–(44), then, similar to (83) and (84), at the decomposition layer of scale j, we have

$$x^{l(j)}_{j,k_j+1} = \frac{\sqrt{2}}{2}\left(x^{\frac{l(j)}{2}}_{j-1,k_{j-1}+1} + x^{\frac{l(j)}{2}}_{j-1,k_{j-1}+2}\right) \tag{89}$$

$$x^{l(j)+1}_{j,k_j+1} = \frac{\sqrt{2}}{2}\left(x^{\frac{l(j)}{2}}_{j-1,k_{j-1}+1} - x^{\frac{l(j)}{2}}_{j-1,k_{j-1}+2}\right) \tag{90}$$

and $y^{l(j)}_{j,k_j}, y^{l(j)+1}_{j,k_j}, z^{l(j)+1}_{j,k_j}$ and $z^{l(j)+1}_{j,k_j}$ can be obtained as in (85)–(88). By using the linear property of expectation and Assumption 1, we can obtain the corresponding expectations and covariances.

Appendix B. Proof of Lemma 2

Denote

$$\mathcal{E}^{l(j)}(k_j) \triangleq \left[e^{l(j),T}_{j,k_j-2^{J-j}+1}, e^{l(j),T}_{j,k_j-2^{J-j}+2}, \cdots, e^{l(j),T}_{j,k_j}\right]^T \tag{91}$$

From (28), we have

$$\begin{aligned}\sum_{k_0=0}^{N}\|e^0_{0,k_0}\|^2 = \|\mathcal{E}^0(k_0=2^J-1)\|^2 + \cdots + \|\mathcal{E}^0(k_0=M2^J-1)\|^2 = \left\|T^{J|0,T}\begin{bmatrix}\mathcal{E}^0(k_J=0)\\ \mathcal{E}^1(k_J=0)\\ \vdots \\ \mathcal{E}^{2^J-1}(k_1=2^{J-1}-1)\end{bmatrix}\right\|^2 + \cdots \\ + \left\|T^{J|0,T}\begin{bmatrix}\mathcal{E}^0(k_J=M)\\ \mathcal{E}^1(k_J=M)\\ \vdots \\ \mathcal{E}^{2^J-1}(k_1=M2^{J-1}-1)\end{bmatrix}\right\|^2 = \sum_{l(J)}\sum_{s=l(J)}^{l(J)+1}\sum_{k_J=0}^{M-1}\|e^s_{J,k_J}\|^2\end{aligned} \tag{92}$$

Appendix C. Proof of Theorem 3

It can be known from Theorem 2 that u^0_{0,k_0} can minimize the cost functions (47) and (48), then from Lemma 2, we have

$$\lambda_{\min}E\left\{\frac{1}{2}e^T_{0,N}Fe_{0,N} + \frac{1}{2}\sum_{k_0=0}^{N-1}e^T_{0,k_0}Qe_{0,k_0}\right\} \leqslant \lambda_{\min}\lambda_{\max}E\left\{\frac{1}{2}\sum_{l(J)}\sum_{s=l(J)}^{l(J)+1}\sum_{k_J=0}^{M-1}\|e^s_{J,k_J}\|^2\right\} \leqslant \lambda_{\max}\sum_{l(J)}\sum_{s=l(J)}^{l(J)+1}J^{s,*} \tag{93}$$

Thus, which yields (74).

References

[1] B.D.O. Anderson, J.B. Moore, Optimal Control: Linear Quadratic Methods, Prentice-Hall, Englewood Cliffs, NJ, 1990.
[2] M. Athans, The role and use of the stochastic linear-quadratic-Gaussian problem in control system design, IEEE Trans. Autom. Control AC-16 (1971) 529–552.
[3] A.N. Akansu, R.A. Haddad, Multiresolution Signal Decomposition: Transforms, Subbands, and Wavelets, Academic Press, Boston, MA, 1992.
[4] C. Bruni, D. Iacoviello, Some results about the optimal LQG tracking problem, Int. J. Control 74 (1999) 977–987.
[5] M. Basseville, A. Benveniste, K. Chou, S. Golden, R. Nikoukhah, A.S. Willsky, Modeling and estimation of multiresolution stochastic processes, IEEE Trans. Inf. Theory 38 (1992) 766–784.
[6] G. Bhatnagar, Q.M. Jonathan Wu, Ba. Raman, Discrete fractional wavelet transform and its application to multiple encryption, Inf. Sci. 223 (2013) 297–316.
[7] Y. Bodyanskiy, O. Vynokurova, Hybrid adaptive wavelet-neuro-fuzzy system for chaotic time series identification, Inf. Sci. 220 (2013) 170–179.
[8] T. Cimen, S.P. Banks, Nonlinear optimal tracking control with application to super-tankers for autopilot design, Automatica 40 (2004) 1845–1863.
[9] M. Chadli, S. Aouaouda, H.R. Karimi, P. Shi, Robust fault tolerant tracking controller design for a VTOL aircraft, J. Franklin Inst. 350 (2013) 2627–2645.
[10] M.Y. Cui, Z.J. Wu, X.J. Xie, Stochastic modeling and tracking control for a two-Link planar rigid robot manipulator, Int. J. Innov. Comput. Inf. Control 9 (2013) 1769–1780.
[11] K. Chou, A.S. Willsky, A. Benveniste, Multiscale recursive estimation, data fusion, and regularization, IEEE Trans. Autom. Control 39 (1994) 464–478.
[12] K. Chou, A.S. Willsky, R. Nikoukhah, Multiscale systems, Kalman filters, and Riccati equations, IEEE Trans. Autom. Control 39 (1994) 479–492.
[13] R.R. Coifman, M.V. Wickerhauser, Entropy-based algorithms for best basis selection, IEEE Trans. Inf. Theory 38 (1992) 713–718.
[14] A. Domahidi, A.U. Zgraggen, M.N. Zeilinger, M. Morari, C.N. Jones, Efficient interior point methods for multistage problems arising in receding horizon control, in: Proceeding of the 51st IEEE Conference on Decision and Control, Hawaii, USA, 2012, pp. 668–674.
[15] H.J. Gao, T.W. Chen, Network-based H_∞ output tracking control, IEEE Trans. Autom. Control 53 (2008) 655–667.
[16] S.S. Hu, Z.Q. Wang, W.L. Hu, Optimal Control Theory and Systems, Science Press, Beijing, 2005.
[17] L. Hong, Multiresolutional distributed filtering, IEEE Trans. Autom. Control 39 (1994) 853–856.
[18] L. Hong, D. Wicker, A spatial-domain multiresolutional particle filter with thresholded wavelets, Signal Process. 87 (2007) 1384–1401.
[19] X.H. Han, X.M. Chang, An intelligent noise reduction method for chaotic signals based on genetic algorithms and lifting wavelet transforms, Inf. Sci. 218 (2013) 103–118.
[20] G. Arantes Jr., L.S. Martins-FilhoL, A.C. Santana, Optimal on-off attitude control for the Brazilian multimission platform satellite, Math. Probl. Eng. (2009) 750945.

[21] A.E.B. Lim, J.B. Moore, L. Faybusovich, Separation theorem for linearly constrained LQG optimal control, Syst. Control Lett. 28 (1996) 227–235.
[22] D. Li, F.C. Qian, P.L. Fu, Optimal nominal dual control for discrete-time linear-quadratic Gaussian problems with unknown parameters, Automatica 44 (2008) 119–127.
[23] S. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, IEEE Trans. Pattern Anal. Mach. Intell. 11 (1989) 674–693.
[24] Y.H. Peng, Wavelet Transform and Engineering Application, Science Press, Beijing, 1999.
[25] O. Rioul, M. Vetterli, Wavelets and signal processing, IEEE Signal Proc. Mag. 8 (1991) 14–38.
[26] M.S. Reis, A multiscale empirical modeling framework for system identification, J. Process Control 19 (2009) 1546–1557.
[27] Q.K. Shen, B. Jiang, P. Shi, J. Zhao, Cooperative adaptive fuzzy tracking control for networked unknown nonlinear multiagent systems with time-varying actuator faults, IEEE Trans. Fuzzy Syst. 22 (2013) 494–504.
[28] C.S. Sastry, S. Rawat, A.K. Pujari, V.P. Gulati, Network traffic analysis using singular value decomposition and multiscale transforms, Inf. Sci. 177 (2007) 5275–5291.
[29] G. Stephanopoulos, O. Karsligil, S.M. Dyer, Multiscale theory for linear dynamic processes part 1. Foundations, Comput. Chem. Eng. 32 (2008) 857–884.
[30] G. Stephanopoulos, O. Karsligil, S.M. Dyer, Multiscale theory for linear dynamic processes part 2. Multiscale model predictive control (MS-MPC), Comput. Chem. Eng. 32 (2008) 885–912.
[31] G.Y. Tang, H.Y. Sun, Y.M. Liu, Optimal tracking control for discrete time-delay systems with persistent disturbances, Asian J. Control 8 (2006) 135–140.
[32] S. Tedmori, N. Al-Najdawi, Image cryptographic algorithm based on the Haar wavelet transform, Inf. Sci. 269 (2014) 21–34.
[33] H.Q. Wang, B. Chen, C. Lin, Adaptive neural tracking control for a class of stochastic nonlinear systems with unknown dead-zone, Int. J. Innov. Comput. Inf. Control 9 (2013) 3257–3269.
[34] Y. Wang, S. Boyd, Fast model predictive control using online optimization, IEEE Trans. Control Syst. Technol. 18 (2010) 267–278.
[35] M.V. Wickerhauser, INRIA lectures on wavelet packet algorithms, in: INRIA Lectures on Wavelet Packet Algorithms, 1991, pp. 31–99.
[36] D.Z. Xu, B. Jiang, P. Shi, Global robust tracking control of non-affine nonlinear systems with application to yaw control of UAV helicopter, Int. J. Control Autom. Syst. 11 (2013) 957–965.
[37] K. You, L. Xie, Linear quadratic Gaussian control with quantised innovations Kalman filter over a symmetric channel, IET Control Theory Appl. 5 (2011) 437–446.
[38] M.G. Yoon, V.A. Ugrinovskii, Robust tracking problem for continuous time stochastic uncertain systems, in: Proceeding of the 42nd IEEE Conference on Decision and Control, Hawaii, USA, 2003, pp. 282–287.
[39] Y.Y. Yin, P. Shi, F. Liu, Gain-scheduled PI tracking control on stochastic nonlinear systems with partially known transition probabilities, J. Frank. Inst. 348 (2011) 685–702.
[40] Q. Zhou, P. Shi, J.J. Lu, S.Y. Xu, Adaptive output-feedback fuzzy tracking control for a class of nonlinear systems, IEEE Trans. Fuzzy Syst. 19 (2011) 972–982.
[41] J.K. Zhou, Q.L. Hu, G.F. Ma, Decentralized adaptive output feedback attitude synchronization tracking control of satellite formation flying, Int. J. Innov. Comput. Inf. Control 8 (2012) 977–988.
[42] L. Zhang, X.L. Wu, Q. Pan, H.C. Zhang, Multiresolution modeling and estimation of multisensor data, IEEE Trans. Signal Process. 52 (2004) 3170–3182.
[43] L. Zhao, Y.M. Jia, Robust transcale state estimation for multiresolution discrete-time systems based on wavelet transform, IET Sig. Process. 7 (2013) 228–238.
[44] L. Zhao, Y. Lin, Y.M. Jia, Y.W. Li, J. Zhang, Transcale LQ tracking control for a class of discrete systems based on wavelet packet decomposition, in: Proceeding of the 2013 American Control Conference, Washington, DC, USA, 2013, pp. 378–383.
[45] L. Zhao, Y.M. Jia, Transcale LQG tracking control for a class of discrete stochastic systems, Eng. Appl. Artif. Intell. 30 (2014) 129–136.