KF is less than 3 % with the 9×9 partition (total of 69 domains) and 101 members. In the case of a 40 m telescope with a 80×80 SH-WFS (and the same r_0), by keeping local domains with the same maximum number of actuators, we can reach a similar coherent energy loss with the same number of members (m = 101). For the 16 m telescope with the 9×9 partition, there is no more than 16 actuators per domain. For the 40 m telescope, in order to have no more than 16 actuators per domain, we must take a 21×21 partition (total of 373 domains). In this AO configuration, our numerical simulations have confirmed that the loss of coherent energy is indeed less than 3%. Each cycle of the AO loop has 3 steps: all the local updates (calculated independently on each local domain), the prediction and the calculation of voltages u_k.

Therefore, we propose to consider a multi-core architecture (single computer or computation cluster), where each core is assigned to only one local domain and performs the local update computations. By using the results of Table 3 and Table 4 in Appendix B, Table 1 gives the theoretical numbers of floating-point operations during one cycle of the AO loop.

Table 1. Numbers of operations with m = 101 during one cycle of the AO loop.

D (m)	n	p	Partition	n_{max}	p_{max}	Update (1 domain)	Prediction	u_k
16	1754	1624	9×9	32	128	7.65×10^6	0.71×10^6	0.76×10^6
40	10370	10048	21×21	32	128	7.65×10^6	4.19×10^6	6.56×10^6

Table 2 gives real speed-tests with a first version of our OpenMP parallelized code: we used a workstation with two Intel(R) Xeon(R) E5-2680 v2 CPUs (total of 20 cores). We have measured the time required for calculating, the update step for one local domain by using only one core, the prediction step by using all the 20 cores and the voltages u_k by using also all the 20 cores.

Table 2. Runtime (in msec) with m = 101 during one cycle of the AO loop.

D (m)	Partition	Update for 1 domain (using 1 core)	Prediction (using 20 cores)	u_k (using 20 cores)	Total
16	9×9	0.94	0.14	0.1	1.18
40	21×21	0.94	0.7	1	2.64

In the AO configuration on a 16 m telescope, we have only 69 domains. Nowadays we can have a single computer with 120 cores (8×15-cores CPUs) similar to the cores in our test workstation. Table 2 shows that, with this kind of single computer (which has more cores than the number of requested domains), our algorithm needs a total of 1.18 ms for one cycle of the AO loop. Therefore, we can easely implement and use our algorithm with a 500 Hz frequency for the AO loop. For the AO configuration on a 40 m telescope, the situation is not basically different: there are many more local domains, but during the update step, computations on each local domain take exactly the same time as in the 16 m case. The time required for the prediction step and the u_k calculation became not negligible: however both of these computations (prediction estimate $\bar{x}_{k+1/k}$ and voltages u_k) are rather well scalable with a larger number of cores. Right now, we cannot have a single computer with so many cores (373), and we have to consider computation cluster which implies additional time for communication stage between the update step and the prediction step. On our computation cluster with infiniBand 4x QDR network, this communication stage takes ~ 2 msec. However the last generation of infiniBand (12x EDR) is claimed to be almost 10 times faster, which makes times for communication reasonable small and a real implementation on a 40 m telescope achievable.

#212876 - $15.00 USD Received 30 May 2014; revised 18 Jul 2014; accepted 29 Jul 2014; published 21 Aug 2014

6. Conclusion

We have presented a new control law, the Local ETKF, based on the LQG approach, a KF adaptation with localizations for large-scale AO systems and the assumption of a perfect DM dynamics. The advantages of this proposed method are significant. Firstly, as all the local Kalman gains can be calculated with empirical covariance matrices at each cycle of the AO loop, it enables to deal with non-stationary behaviors (turbulence, vibrations). Secondly, as the structure of this algorithm is intrinsically parallel, its implementation on ELTs can easily be done on a parallel architecture (CPUs/GPUs cluster) with a reduced computational cost: let us remind that the complexity of the update step does not depend on the diameter of the telescope but only on the maximum number of actuators per each local domain, which significantly speeds up the algorithm. The more subpartitions in the pupil of the telescope, the better is the performance, both in terms of coherent energy and of runtime. In our simulations with a Von Karman turbulence and a SCAO system with an adequate partition of the pupil, the loss of coherent energy compared with the optimal solution given by the KF can be less than 3 %. We have presented numerical simulations in the case of a SCAO system with an AR1 turbulence model on a zonal basis, but we can already extend this method with an AR2 model. For the runtime tests we have already developed an OpenMP parallelized version, but we are also currently working on a new version with GPUs on the COMPASS platform [41].

In the short-term, there are different points to study more in details. The first one is the influence of the turbulence charateristics on the size of the local observation regions which reflects the distance, called the localization length, over which the correlations calculated with the ensemble's members are not meaningful. In the same way, we need to evaluate the influence of the spider arms on the choice of the subpartitions on the pupil of the telescope. In order to resolve the problem of differential pistons, we have to improve our fast least-squares based method by taking into account the error covariance matrices. Then, the robustness of the Local ETKF must be also studied when the parameters of the turbulence change. We must indeed estimate the impact of the turbulence model error (wind speed, turbulence profile, L_0), first in absence of turbulence model update. Then, assuming some turbulence model identification and update, one shall consider the gain brought by this non-stationary control solution, its stability and the speed of convergence.

In the long-term, we have of course to demonstrate the potentials of the Local ETKF in the case of wide field tomographic AO and to consider the extension to limited DM dynamics. Afterwards, two other important aspects already developed in geophysics must be also explored in AO. The first one is when the operator C in Eq. (6) is non-linear [42]. The EnKF-based method does not require a linearized model, an advantage over the KF-based method, and this could be very suitable for non-linear WFS. The second one is the possibility of asynchronous observations assimilation [37, 43], for the multi-rate case in the prospect of Natural Guide Star and Laser Guide Star wavefront sensing for wide field tomographic AO.

Appendix A: Mathematical expression of the update estimate $\hat{x}_{k/k}$ for the ETKF

For computational reasons, it is better to change the original expression of the update estimate $\hat{x}_{k/k}$ in Eq. (9) by using the following version of the Sherman-Morrison-Woodbury identity:

$$(U \times U^{\mathrm{T}} + S \times S^{\mathrm{T}})^{-1} = (S^{-1})^{\mathrm{T}}\{S^{-1} - (S^{-1}U)[\mathrm{Id} + (S^{-1}U)^{\mathrm{T}}(S^{-1}U)]^{-1}(S^{-1}U)^{\mathrm{T}}S^{-1}\}. \quad (28)$$

By replacing the two expressions from Eqs. (11) and (13) in Eq. (9), we obtain:

$$\hat{x}_{k/k} = \bar{x}_{k/k-1} + Z_{k/k-1}Z_{k/k-1}^{\mathrm{T}}\mathrm{C}_1^{\mathrm{T}}(\mathrm{C}_1 Z_{k/k-1}Z_{k/k-1}^{\mathrm{T}}\mathrm{C}_1^{\mathrm{T}} + \Sigma_w)^{-1}(y_k - \bar{y}_{k/k-1}). \quad (29)$$

#212876 - $15.00 USD Received 30 May 2014; revised 18 Jul 2014; accepted 29 Jul 2014; published 21 Aug 2014

We can note $U = \mathrm{C}_1 Z_{k/k-1}$ and as Σ_w is a strictly positive diagonal matrix, we can note $S = \Sigma_w^{1/2}$ which is very easy to compute and to invert. By using the notations U and S, we can identify $(\mathrm{C}_1 Z_{k/k-1} Z_{k/k-1}^{\mathrm{T}} \mathrm{C}_1^{\mathrm{T}} + \Sigma_w)^{-1}$ as the left-hand side of Eq. (28) and by using the right-hand side of Eq. (28) in Eq. (29), we can obtain a new expression for $\hat{x}_{k/k}$:

$$\begin{aligned}\hat{x}_{k/k} = \bar{x}_{k/k-1} + Z_{k/k-1}(\Sigma_w^{-1/2}\mathrm{C}_1 Z_{k/k-1})^{\mathrm{T}}\{\Sigma_w^{-1/2}(y_k - \bar{y}_{k/k-1}) - \Sigma_w^{-1/2}\mathrm{C}_1 Z_{k/k-1} \times \\ [\mathrm{Id} + (\Sigma_w^{-1/2}\mathrm{C}_1 Z_{k/k-1})^{\mathrm{T}}(\Sigma_w^{-1/2}\mathrm{C}_1 Z_{k/k-1})]^{-1}(\Sigma_w^{-1/2}\mathrm{C}_1 Z_{k/k-1})^{\mathrm{T}}\Sigma_w^{-1/2}(y_k - \bar{y}_{k/k-1})\}.\end{aligned} \tag{30}$$

Let us define the vector $S_{inov} = \Sigma_w^{-1/2}(y_k - \bar{y}_{k/k-1})$ and the matrix $S_{cz} = \Sigma_w^{-1/2}\mathrm{C}_1 Z_{k/k-1}$, then Eq. (30) becomes:

$$\hat{x}_{k/k} = \bar{x}_{k/k-1} + Z_{k/k-1} S_{cz}^{\mathrm{T}}\{S_{inov} - S_{cz}[\mathrm{Id} + S_{cz}^{\mathrm{T}} S_{cz}]^{-1} S_{cz}^{\mathrm{T}} S_{inov}\}. \tag{31}$$

By using the EVD of the matrix $(\mathrm{Id} + S_{cz}^{\mathrm{T}} S_{cz})$ (which gives the expression of the ETM T_k in Eq. (16)), we obtain a new decomposition for the matrix inversion in the square brackets of last Eq. (31): therefore, it does not require the inversion of the $p \times p$ matrix in the brackets of the original Eq. (11), but only a $m \times m$ matrix EVD which can be computationally cheaper if $m \ll p$. We finally obtain for the update estimate:

$$\hat{x}_{k/k} = \bar{x}_{k/k-1} + Z_{k/k-1} S_{cz}^{\mathrm{T}}\{S_{inov} - S_{cz} Q_k \Gamma_k^{-1} Q_k^{\mathrm{T}} S_{cz}^{\mathrm{T}} S_{inov}\}. \tag{32}$$

Appendix B: Total number of operations for the ETKF on a zonal basis

The mathematical formalism presented in this paper is valid for both a modal or a zonal basis. But using a zonal basis (where the phase is estimated on the locations of each valid actuator of the DM) enables to compute some very sparse matrices which is very suitable for calculations on large-scale AO systems. Let us define $\mathrm{n_{act}}$ the number of valid actuators: on a zonal basis with an AR1 turbulence model in the ETKF-based control law, as $\mathrm{A_{tur}} = \mathrm{a_{tur}} \times \mathrm{Id}$, the extracted matrix $\mathrm{A}_1 = \begin{pmatrix} \mathrm{A_{tur}} & 0 \\ \mathrm{Id} & 0 \end{pmatrix}$ is composed by only two $\mathrm{n_{act}} \times \mathrm{n_{act}}$ diagonal blocks. Thus, A_1 is a very sparse matrix and its size is $\mathrm{n} \times \mathrm{n}$, where $\mathrm{n} = 2 \times \mathrm{n_{act}}$. Let us define $\mathrm{p_{sap}}$ the number of valid subapertures of the S-H WFS (each subaperture gives two slopes of the residual phase in two directions): the dimension of the measurement vector y_k is $\mathrm{p} \times 1$, where $\mathrm{p} = 2 \times \mathrm{p_{sap}}$. Let us define the matrix D_1, modeling the S-H WFS on the zonal basis: this matrix D_1 is the equivalent of the linear operator D in Eq. (3). For a SH-WFS with a Fried geometry, D_1 enables to calculate 2 slopes (at the center of each subaperture) from the 4 estimations of the turbulent phase on the actuators at the 4 corners of each subaperture: its size is $\mathrm{p} \times \mathrm{n_{act}}$ and each row of this matrix is then composed by only 4 non-zero values. Thus, D_1 is a very sparse matrix. Moreover the extracted observation matrix $\mathrm{C}_1 = \begin{bmatrix} 0 & \mathrm{D}_1 \end{bmatrix}$ is also very sparse and its size is $\mathrm{p} \times \mathrm{n}$. The influence matrix N characterises the DM on the zonal basis by Eq. (4): its size is $\mathrm{n_{act}} \times \mathrm{n_{act}}$ and it is a very sparse matrix. For the calculation of $\bar{y}_{k/k-1}$, we have to compute $\mathrm{D}_1\mathrm{N} \times u_{k-2}$ where the matrix $\mathrm{D}_1\mathrm{N}$ is the result of the mulplication of 2 sparse matrices. In our classical AO configuration, for a 16 m, a 32 m and a 40 m telescope, the number of non-zero values per row of this matrix $\mathrm{D}_1\mathrm{N}$ is always less than 36.

B.1 The Update step

By using the last expression (Eq. (32)) and the associativity property of matrix multiplication, we can compute many matrix-vector multiplications in order to minimize as far as possible the theoretical numerical cost.

#212876 - $15.00 USD Received 30 May 2014; revised 18 Jul 2014; accepted 29 Jul 2014; published 21 Aug 2014

Actually the significant computational expense is the EVD and the 3 matrix-matrix multiplications: $S_{cz}^{\mathrm{T}} \times S_{cz}$, $Q_k \times (\sqrt{m-1}\Gamma_k^{-1/2} Q_k^{\mathrm{T}})$ and $Z_{k/k-1} \times (Q_k \sqrt{m-1}\Gamma_k^{-1/2} Q_k^{\mathrm{T}})$.

Table 3. Numbers of multiplications during the update step.

Expressions	Multiplications	Result Size
$\bar{y}_{k/k-1} = \mathrm{C}_1 \times \bar{x}_{k/k-1} - \mathrm{D}_1\mathrm{N} \times u_{k-2}$	$p \times 4 + p \times 36$	(p,1)
$S_{inov} = \Sigma_w^{-1/2} \times (y_k - \bar{y}_{k/k-1})$	p	(p,1)
$S_{cz} = \Sigma_w^{-1/2}\mathrm{C}_1 \times Z_{k/k-1}$	$p \times 4 \times m$	(p,m)
$S_{cz}^{\mathrm{T}} \times S_{cz}$	$m \times p \times m$	(m,m)
EVD of $(\mathrm{Id} + S_{cz}^{\mathrm{T}} S_{cz})$	m^3	(m,m)
Γ_k^{-1} and $\Gamma_k^{-1/2}$	$m + \gamma \times m$	(m,m)
$S_{cz}^{\mathrm{T}} \times S_{inov}$	$m \times p$	(m,1)
$S_{cz} \times (Q_k \times (\Gamma_k^{-1} \times (Q_k^{\mathrm{T}} \times (S_{cz}^{\mathrm{T}} S_{inov}))))$	$m^2 + m + m^2 + p \times m$	(p,1)
$S_{cz}^{\mathrm{T}} \times (S_{inov} - S_{cz} Q_k \Gamma_k^{-1} Q_k^{\mathrm{T}} S_{cz}^{\mathrm{T}} S_{inov})$	$m \times p$	(m,1)
$Z_{k/k-1} \times S_{cz}^{\mathrm{T}}(S_{inov} - S_{cz} Q_k \Gamma_k^{-1} Q_k^{\mathrm{T}} S_{cz}^{\mathrm{T}} S_{inov})$	$n \times m$	(n,1)
$Z_{k/k-1} \times (Q_k \times ((\sqrt{m-1} \times \Gamma_k^{-1/2}) \times Q_k^{\mathrm{T}}))$	$m + m^2 + m^3 + n \times m \times m$	(n,m)

Total number of multiplications: $(m^2+m) \times n + (m^2+7m+41) \times p + 2m^3 + 3m^2 + (3+\gamma)m$. The number of additions is the same order of magnitude as the number of multiplications.

B.2 The Prediction step

By using a zonal basis with an AR1 turbulence model, A_1 is composed by two $n_{act} \times n_{act}$ diagonal blocks, one of them is the identity matrix. Therefore, multiplying A_1 with a vector is reduced to only one multiplication with the block $\mathrm{A}_{\mathrm{tur}}$ (the other one consists on a copy).

Table 4. Numbers of operations during the prediction step.

Expressions	Multiplications	Additions	Result Size
$X_{k+1/k} = \mathrm{A}_1 \times X_{k/k} + V_{k+1}$	$\frac{n}{2} \times m$	$\frac{n}{2} \times m$	(n,m)
$\bar{x}_{k+1/k} = \frac{1}{m}\sum_{i=1}^{m} x_{k+1/k}^i$	n	$n \times (m-1)$	(n,1)
$Z_{k+1/k} = \frac{[x_{k+1/k}^1 - \bar{x}_{k+1/k};\ldots;x_{k+1/k}^m - \bar{x}_{k+1/k}]}{\sqrt{m-1}}$	$n \times m$	$n \times m$	(n,m)

Total number: $(\frac{3}{2}m+1) \times n$ multiplications + $(\frac{5}{2}m-1) \times n$ additions.

B.3 The Projection onto the DM

The voltage u_k is calculated with the MVM: $u_k = (\mathrm{N}^{\mathrm{T}}\mathrm{N})^{-1}\mathrm{N}^{\mathrm{T}} \times \hat{\varphi}_{k+1/k}^{\mathrm{tur}}$.

The matrix $\mathrm{P} = (\mathrm{N}^{\mathrm{T}}\mathrm{N})^{-1}\mathrm{N}^{\mathrm{T}}$ is a sparse matrix on a zonal basis. In our SCAO system, for a 16 m (respectively a 40 m) telescope, the sparsity of this matrix is 51 % (respectively 88 %). Actually, on a zonal basis, the matrix P can be much more sparse when, for each actuator, we take into account only the neighboring actuators close to less than 2 pitches.

Acknowledgments

This work has been supported with financial grants from the cross-disciplinary mission MASTODONS of the CNRS and from the Programme Hubert Curien-AURORA (Campus France) for mobilities between France and Norway.

#212876 - $15.00 USD Received 30 May 2014; revised 18 Jul 2014; accepted 29 Jul 2014; published 21 Aug 2014