volume average, $d_r^i * = N_r^{-1}(d_r^i - \bar{d}_r)$ in Eq. (A7). The SVD of the matrix of normalized residuals (Eq. (A8)) is evaluated, which yields the tracer variability vertical EOFs and coefficients, $\mathbf{E}_{trc}^z = \mathbf{N}_r \mathbf{E}_r^*$ and $\Pi_{trc}^z = \Sigma_r^{*2}/s$, or dominant decomposition of $\mathbf{C}_{trc}^z$ (Eq. (A9)). Finally, including the scalar horizontal variances of T and S into $C_{trc}^z \epsilon \mathbb{R}^{2l_v \times 2l_v}$, the 3D covariance matrix $\mathbf{P}_{trc} \in \mathbb{R}^{2l \times 2l}$ is simply the Kronecker product $\mathbf{C}_{trc}^z \otimes \mathbf{C}_{trc}^{\mathbf{r}*}$ (Eq. (A10)). The significant rank-p eigendecomposition of $\mathbf{P}_{trc}$, $\mathbf{E}_{trc} \Pi_{trc} \mathbf{E}_{trc}^T$ where $\mathbf{E}_{trc} \epsilon \mathbb{R}^{2l \times q}$ (Eq. (A11)), is then easily obtained from the eigendecompositions of $\mathbf{C}_{trc}^z$ and $\mathbf{C}_{trc}^{\mathbf{r}*}$ (Graham, 1981).

In the second stage, an ensemble of q tracer initial conditions, $\hat{\psi}_{trc}^j$, of covariance matrix (Eq. (A11)) is first created (Eq. (A12)). The fields $\hat{\psi}_{trc}^j$ are obtained by adding to $\hat{\psi}_{trc}$ an adequately weighted eigenvector j, $\mathbf{E}_{trc} \Pi_{trc}^{1/2} \sqrt{q}\, \mathbf{e}^j$, where the $\mathbf{e}^j$'s are $j = 1, \ldots, q$ base vectors. The resulting states are then balanced by an ensemble of adjustment PE integrations: the perturbed tracer fields $\hat{\psi}_{trc}^j$ are fixed and the momentum equations in Eq. (A1) integrated forward until the mean kinetic energy stabilizes around a plateau, without rapid changes (parallel computing is then used). The differences between Ψ_0 and these PE adjusted fields $\hat{\psi}_0^j$ form the matrix $\mathbf{M}$ (Eq. (A13)). This matrix is normalized and the initial ES (Eq. (A15)) is estimated from the SVD of $\mathbf{M}^*$ (Eq. (A14)). During the parallel adjustment PE integrations, a similarity coefficient is evaluated (as in the ES forecast, Table 5 hereafter) to assess the added value of new integrations and thus decide at which j to stop. Finally, the factor γ^2 in Eq. (A15) scales the variability variance to an error variance.

A.2. Assimilation or data-forecast melding

The melding is chosen linear and based on a minimum error variance in the sample ES (Table 6). The present scheme being recursive, the sample ES forecast, described by $\mathbf{E}_-$ and $\Pi(-)$, is assumed available. It is obtained in Section A.4.

The update of the state (Eq. (A16)) uses the gain $\mathbf{K}^p$ (Eq. (A17)), optimal for the dominant error covariance forecast given by $\mathbf{E}_-$ and $\Pi(-)$. In Eq. (A18), $\Pi(+)$ is estimated by eigendecomposition of the right-hand-side: the columns of $\mathbf{H}$ are ordered orthonormal eigenvectors and $\Pi(+)$ is the ordered diagonal matrix of eigenvalues; $\mathbf{E}_+$ then follows from Eq. (A19).

A.3. Adaptive learning of the dominant errors

For several reasons, including the simple dynamical and measurement error models, the error subspace reduction and the linear melding, significant components of the ocean signal could be left over in the a posteriori residuals. Table 7 describes the discrete algorithm used here to learn (correct) the ES in accord with these possibly significant residuals. Continuous dynamical systems (e.g., Brockett, 1990) for such adaptive ES learning and ESSE assimilations can also be derived (Lermusiaux, 1997).

The a posteriori tracer residuals, $\mathbf{d} - \mathbf{C}\hat{\psi}(+)$, are first analyzed into gridded fields $\hat{n}(+)$, via a one-stage ESSE analysis (Eqs. (A20) and (A21)). The background for these