Assume a high-dimensional CT discretization $x(t)$ of an infinite-dimensional system (say, a flow of air and smoke over a domain $\Omega$, the many important details of which will be omitted completely here in order to focus solely on the larger questions of state estimation and adaptive observation of such a system), excited by state disturbances $w(t)$ [CT, white, with zero mean and spectral density $Q(t)$], governed by

$$\frac{dx(t)}{dt} = f(x(t), w(t)).$$

(1)

Assume also that there are $M$ sensor vehicles ($i = 1, \ldots, M$) moving (in CT) throughout $\Omega$, with positions $q_i(t)$ and control inputs $u(t)$, and taking measurements $y_k$ (in DT, at times $t_k = kh$) corrupted by measurement noise $\nu_k$ [DT, white, Gaussian, with zero mean and covariance $R_k$], according to:

$$\frac{dq_i(t)}{dt} = g(q_i(t), u(t)),
\begin{align*}
y_k &= h_k(x_k, q_k) + \nu_k = H_k(q_k) x_k + \nu_k.
\end{align*}

(2)

That is, we assume the (DT) measurements $h_k(x_k, q_k)$ are actually linear functions of the state $x(t_k) = x_k$, but the measurement matrix $H_k(q_k)$ corresponding to the $i$th sensor vehicle is a function of its position at time $t_k$, denoted $q(t_k) = q_k$ [which, in turn, can be changed by modifying the control inputs $u(t)$ in (2)]. It is assumed for the purpose of this derivation that the CT trajectories of the sensor vehicles can be modeled accurately, and thus a disturbance input to (2) is not included in the model, and an estimator for the vehicle positions $q(t)$ is not needed.

The Ensemble Kalman filter proceeds as follows: individual ensemble members $\hat{x}_i^j(t)$ evolve according to the modeled state equation (1) and are excited by disturbances $w^j(t)$, and individual measurements $y_k$ [taken according to (2)-(3)] are corrupted with noise $\nu_k^j$ in each ensemble update, with statistics consistent with the models of the random processes $w(t)$ and $\nu_k$ (i.e., with spectral density of $Q(t)$ and covariance $R_k$, respectively, and zero mean):

$$\begin{align*}
\text{for } j = 1, \ldots, N & : \quad \frac{d\hat{x}_i^j(t)}{dt} = f(\hat{x}_i^j(t), w_i^j(t)), \quad \hat{d}_i^j = x_i + \nu_i^j,
\text{with } \hat{x}_i^j[k] = \hat{x}_i^j[t(k)] \text{ before the update, } \hat{x}_i^j[k] = \hat{x}_i^j(t(k)) \text{ after the update,}
\end{align*}$$

(4)

where the (low-rank) ensemble approximation $P^e(t)$ of the covariance $P(t) = \mathbb{E}\{[x(t) - \hat{x}(t)][x(t) - \hat{x}(t)]^T\}$ is

$$P^e(t) = \frac{(\delta X)(\delta X)^T}{N - 1}, \quad \delta X = \begin{bmatrix} \delta \hat{x}_1^1 & \delta \hat{x}_2^1 & \cdots & \delta \hat{x}_N^1 \\
\delta \hat{x}_1^2 & \delta \hat{x}_2^2 & \cdots & \delta \hat{x}_N^2 \\
\vdots & \ddots & \ddots & \vdots \\
\delta \hat{x}_1^N & \delta \hat{x}_2^N & \cdots & \delta \hat{x}_N^N \end{bmatrix},$$

(5)

$$\delta \hat{x}_i^j(t) = \hat{x}_i^j(t) - \hat{x}(t), \quad \hat{x}(t) = \frac{1}{N} \sum \hat{x}_i^j(t).$$

(6)

The indexes have been put in distinct locations in order to keep the notation clear (as much as possible); for example, the vector $\nu_k^j$ represents the value of $\nu$ for the $i$th vehicle, the $j$th ensemble member, and the $k$th timestep.

Starting from the current time (taken as $t = 0$ in the discussion that follows), the paper on Adaptive Observation by Zhang et al, which assumed a different model for the evolution of $P$ and used a slightly different (less precise) notation, gave a framework to minimize a cost function $J$ with respect to the control inputs $u(t)$ over the interval $[0, T]$, which we write here in the form

$$J = \text{trace}[P^e(T)] + \frac{1}{2} \sum_{i=1}^M \int_0^T u_i^T(t) Z^i u_i(t) dt.$$
Please do not rewrite the 10 equations above, which set the precise notation to be used, but instead simply refer to them in your writeup by number. Thus, the first numbered equation in your (hand-written) writeup will be (11). You can follow the general approach of section II of Zhang et al (2011) in your writeup, but feel free to modify/extend the presentation style and notation used by Zhang as necessary (but, importantly, stay fully consistent with the notation set in the 10 equations above). This exam is, most likely, unlike any other that you ever have taken. Don’t stress yourself into inaction, but roll with it. You have a solid block of time to do this; I recommend that you work through all the equations first, in rough draft form (that you do not turn in), then rewrite your entire derivation neatly, including all useful connecting explanatory phrases, to tie it all together clearly, like in a research paper. You will be given an overall grade for your final presentation related to both its accuracy and clarity, akin to that of a paper review (i.e., not 2 points for this, 3 points for that...). Upon completion of this exam, maybe one or two of you can help me to write the full paper on this topic for your Master’s thesis. If you are interested in doing this, it might be a good idea to take MAE290c next quarter, so we can work together on the (difficult) problem of developing an appropriate formulation (and, simulation code) for the problem described in the first paragraph and stated in (1).