Comments on Notation and Equations (the authors are not obligated to address this):
The polarization calculations in this manuscript are cumbersome and difficult to follow. For example, Eq. 8 has 65 terms (not all unique). Using linear algebra, it should be possible to represent the operation of this instrument in a way that's more intuitive.

Assuming a linear input polarization, it is possible to obtain any *linear* polarization state using a HWP. It is also possible to obtain *all* polarization states on the Poincare sphere with the combination of a HWP and a QWP. This is in accordance with the instrument design. So in the approximation that no other optic in the system is affecting polarization after the wave plates, there is no reason to belabor the definitions of the output polarization or the waveplate Mueller matrices and angles. All the cosines and sines in Eq. 8 are unnecessary. One can easily write the output state of laser A as

$$\mathbf{S}_A = \begin{bmatrix} S_1 & S_2 & S_3 & 0 \end{bmatrix}^T$$

where $S_3 = \sqrt{S_1^2 - S_2^2}$

And laser B as

$$\mathbf{S}_B = \begin{bmatrix} S_1 & S_2 & S_3 & S_4 \end{bmatrix}^T$$

where $S_4 = \sqrt{S_1^2 - S_2^2 - S_3^2}$

As part of the optimization process, we need to determine these vectors. (Note that the elements of these two vectors are not the same and that the degrees of freedom can be further described by the Poincare sphere polar coordinates or normalization).

By a very similar argument, you can fully evaluate the Mueller matrices in the receiver equations with the output vector **e** and describe each channel with a single vector that is the diattenuation vector of the receiver channel. The result is a row vector, analogous to a the transmitted Stokes vector, describing the channel's preference to transmit a particular polarization. Like with the transmitted polarization states, this diattenuation vector can take the form of any *linear* polarization if a linear analyzer is preceded by a HWP. The diattenuation vector can take the form of *any* polarization if a linear analyzer is preceded by a combination HWP and QWP. So now you have the receiver taking the form

$$\mathbf{D}_C = \begin{bmatrix} D_1 & D_2 & D_3 & 0 \end{bmatrix}$$

And

$$\mathbf{D}_D = \begin{bmatrix} D_1 & D_2 & D_3 & D_4 \end{bmatrix}$$

Instead of expanding Mueller matrices of individual polarization elements, why not do what Kaul 2004 and Hayman 2012 did and write that an optical measurement is described by a vector projection of the scattering phase matrix with a measurement described by the combination of the transmitted and detected states?

$$I = \mathbf{p}^T(\mathbf{f} + \mathbf{g})$$

Where **f** and **g** are vectorized scattering matrices.

The intensity measured on the detector is the projection of the scattering matrix via an incident and detected polarization state. The elements of that projection is given by

$$p_{i+4(j-1)} = D_j S_i$$

though there are simplifications that can be made due to redundancies in the scattering matrix as noted in Kaul 2004 and Hayman 2012.

So to optimize the experiment, you need to determine the values of the transmitted Stokes vectors and the receiver diattenuation vectors. The constraints on the polarization states are encompassed by the definitions of those vectors. The exact waveplate angles to achieve those states can be determined afterward. There is no need to write these expanded matrix equations of multiple polarization elements to describe and optimize the measurement.

Also notable, if you vectorize your matrix definitions, you can perform principal component analysis on them to determine the dominant modes in the model study. That also provides a basis for determining the optimal configuration of the system.