Appendix 1

Description of the fully nonlinear Boussinesq wave model FUNWAVE-TVD

The numerical phase resolving model FUNWAVE_TVD is based on the conservative form of the fully nonlinear Boussinesq equations formulated by Shi et al. (2012). Following Tonelli and Petti (2009) wave breaking forces the model to switch from Boussinesq equations, where dispersive and nonlinear effects are of a similar order of magnitude, to the nonlinear shallow water equation, where nonlinearity dominates. This model employs a Total Variation Diminishing (TVD) spatial discretisation scheme to solve the fully non-linear Boussinesq equation (combining finite-volume for nonlinear terms and finite-difference for dispersive terms) and incorporates a time-dependent reference level (Kennedy et al. 2001) moving with the instantaneous free surface to calculate the velocity potential. The combination of the shock-capturing TVD scheme and moving reference provides robust performance in simulating breaking waves and optimising nonlinear behaviour. Furthermore, the model uses an adaptative time stepping defined from a third-order Strong Stability-Preserving (SSP) Runge–Kutta scheme (Gottlieb et al., 2001) to increase model stability. The conservative form of the fully nonlinear Boussinesq equations in FUNWAVE_TVD employs a modification of the leading order pressure term in the momentum equation using a modified surface gradient term such as:

$$\eta_t + \nabla.\mathrm{M} = 0 \tag{A1.1}$$

$$\begin{aligned} M_t + \nabla.\left[\frac{MM}{H_{tot}}\right] + \nabla\left[\frac{1}{2}g(\eta^2 + 2h\eta)\right] \\ = H_{tot}\{\bar{u}_{2,t} + u_\alpha.\nabla\bar{u}_2 + \bar{u}_2.\nabla u_\alpha - V'_{1,t} - V''_1 - V_2 - V_3 - R\} \\ + g\eta\nabla h \end{aligned} \tag{A1.2}$$

where ∇ denotes the horizontal partial derivative $((\partial/\partial x),(\partial/\partial y))$, η is the free surface elevation, h is the water depth, $H_{tot} = h + \eta$ is the total local water depth and g is the gravitational acceleration, the terms $\nabla\left[\frac{1}{2}g(\eta^2 + 2h\eta)\right]$ and $g\eta\nabla h$ are components of the surface gradient. The horizontal volume flux is expressed as:

$$M = H_{tot}\{u_\alpha + \bar{u}_2\} \tag{A1.3}$$

where u_α is the horizontal velocity at the reference level $z_\alpha = \zeta h + \beta\eta$ (from Kennedy et al. (2001)) with ζ=-0.53 and β=0.47. While u_2 is the depth dependant correction at $O(\mu^2)$ (with μ representing the ratio of depth over wave length) that is expressed as:

$$u_2(z) = (z_\alpha - z)\nabla A + \frac{1}{2}(z_\alpha^2 - z^2)\nabla B \tag{A1.4}$$

with $\nabla A = \nabla.(hu_\alpha)$ and $\nabla B = \nabla.u_\alpha$. The depth-averaged contribution to the horizontal velocity field is given by:

$$\bar{u}_2 = \frac{1}{H_{tot}}\int_{-h}^{\eta} u_2(z)dz = \left[\frac{z_\alpha^2}{2} - \frac{1}{6}(h^2| - h\eta + \eta^2)\right]\nabla B + \left[z_\alpha + \frac{1}{2}(h-\eta)\right]\nabla A \tag{A1.5}$$

V_1 and V_2 represent the dispersive terms of the Boussinesq equation defined as:

$$V_1 = \left\{\frac{z_\alpha^2}{2}\nabla B + z_\alpha\nabla A\right\}_t - \nabla\left[\frac{\eta^2}{2}B_t + \eta A_t\right] \tag{A1.6}$$

$$V_2 = \nabla\left\{(z_\alpha - \eta)(U_\alpha.\nabla)A + \frac{1}{2}(z_\alpha^2 - \eta^2)(\mathrm{U}_\alpha.\nabla)B + \frac{1}{2}[A + \eta B]^2\right\} \tag{A1.7}$$

with V_3 representing the second order ($O(\mu^2)$) effect of the vertical velocity, which is expressed as:

$$V_3 = \omega_0 i^z \times \bar{u}_2 + \omega_2 i^z \times u_\alpha \tag{A1.8}$$

Where with i^z the unit vector in the vertical direction and:

$$\omega_0 = (\nabla \times u_\alpha).\, i^z = v_{\alpha,x} - u_{\alpha,y} \tag{A1.9}$$

$$\omega_2 = (\nabla \times \bar{u}_2).\, i^z = z_{\alpha,x}(A_y + z_\alpha B_y) - z_{\alpha,y}(A_x + z_\alpha B_x) \tag{A1.10}$$

R in Eq. A3.2 represents the combination of diffusive (R_s) and dissipative (R_f) terms (Chen et al., 1999) induced by sub-grid lateral turbulent mixing and bottom friction, $R = R_s + R_f$, with R_f, expressed as:

$$R_f = \frac{C_d}{h+\eta} u_\alpha |u_\alpha| \tag{A1.11}$$

where C_d is the bottom friction coefficient.

Appendix 2

Description of the Bispectra Mode Decomposition

Advantages and limitations of orthogonal decomposition methods

Since its first application to fluid mechanics (Lumley, 1967), the Empirical Orthogonal Function (EOF) analysis has been used extensively to identify stationary patterns in random wavefields. However, the limitations of this approach are twofold. First, a single physical wave transformation process can be spread over more than one EOF mode; inversely, more than one physical process can contribute to one EOF mode. Additionally, the orthogonal nature of the EOF modes does not support the complex values necessary to define the physical properties of the wavefield from high-order spectral analysis. Therefore, the EOF analysis cannot establish causality between modal states and physical processes other than physical mechanisms previously accepted in the literature (e.g. standing waves) (Emery and Thomson, 2014).

Development and advantages of high-order spectral decomposition methods

Investigating the generation of stationary patterns from coherent wave amplification requires a decomposition method capable of holding information on both spectral and phase characteristics of the wavefield. Such information can be provided by high-order statistical analyses such as bispectrum, defined from the third moment of the data field (Hasselmann et al., 1963). The bispectrum presents attractive properties to identify coherent wave amplification. It is not only capable of detecting quadratic phase coupling for specific sets of frequencies but also represents a measure of skewness, which is expected to increase in areas of wave ray intersection (e.g. Janssen and Herbers, 2009). Despite these advantages, the bispectrum is only applicable to one-dimensional spatial domains. To overcome this limitation, Schmidt (2020) recently introduced the Bispectra Mode Decomposition (BMD), which consists of maximising the expansion coefficients of a spatial integral measure of the bispectrum. Thus, the BMD can be regarded as a decomposition method based on the same principle as the spectral EOF but applied to higher-order spectral analysis.

Description of the Bispectral Mode decomposition method

In the BMD approach, the time series of two-dimensional sea surface elevation observations defined in the time domain and cartesian coordinate system ($q(\xi,t) \in \mathbb{C}^{M\times N_t}$) are first redefined in the frequency domain using Welch's method (Welch, 1967) such as:

$$\hat{q}(\xi, f_k) = \sum_{j=0}^{N_{FFT}-1} q(\xi, t_{j+1}) e^{-i2\pi jk/} N_{FFT} \quad \text{with } k = 0, \dots, N_{FFT}-1 \tag{A2.1}$$

where $q(\xi, t_j) \in \mathbb{C}^M$ represents the two-dimensional sea surface observations in the spatial domain ξ defined by a number of points $M = N_x, N_y, N_z$ at a sample time t_j with j = 0, …, N_t. N_{FFT} represents the number of samples in one of the N_{blk} segments used to calculate the Fourier transform. Two-dimensional observations are, therefore, redefined in the space-frequency domain $\hat{q}(\xi, f_k) \in \mathbb{C}^{M\times N_{blk}}$.

The product of the Fourier coefficients used to define the bispectrum for frequencies k an l is obtained from the Hadamard product of the matrices $\hat{q}(\xi, f_k) \equiv \hat{q}_k$ and $\hat{q}(\xi, f_l) \equiv \hat{q}_l$ such as:

$$\hat{q}_{k\circ l} = \hat{q}_k \circ \hat{q}_l \tag{A2.2}$$

The spatial integral measure of the bispectrum is therefore expressed as:

$$b(f_k, f_l) = E\left[\int_\Omega \hat{q}_k{}^* \circ \hat{q}_l{}^* \circ \hat{q}_{k+l}\, d\xi\right] = E\left[\hat{q}_{k\circ l}{}^H \hat{q}_{k+l}\right] \tag{A2.3}$$

where $E[.]$ Is the expectation operator, $(.)^*$ and $(.)^H$ denote the complex conjugate and transpose, respectively. Assuming that the observed fluid is incompressible, the form of the triadic interaction in the Navier-Stokes is used in the BMD to establish a causal relationship between the product of the two interacting frequency components represented by the term $\hat{q}_{k\circ l}$ in Eq. A2.3, generating the third frequency component represented by the term $\hat{q}_{k+l}$. Therefore, the interacting and resulting frequency components are linked by a shared expansion coefficient, a_{ij}, in the modal decomposition and defined by the linear expansions:

$$\phi_{k\circ l}^{[i]}(\xi, f_k, f_l) = \sum_{j=1}^{N_{blk}} a_{ij}\,(f_{k+l})\hat{q}_{k\circ l}^{[j]} \tag{A2.4}$$

$$\phi_{k+l}^{[i]}(\xi, f_{k+l}) = \sum_{j=1}^{N_{blk}} a_{ij}\,(f_{k+l})\hat{q}_{k+l}^{[j]} \tag{A2.5}$$

The cross-frequency fields $\phi_{k\circ l}$ are maps of phase alignment between two frequency components, while bispectral modes ϕ_{k+l} are linear combinations of Fourier modes related to the amplitude of oscillations of the sea surface at frequency $k + l$. Consequently, the modal decomposition in the BMD is defined from the spectral properties of each segment obtained from the Welch method rather than from the raw two-dimensional time series of observations conventionally used in the EOF analysis. Eq. A2.4 and A2.4 can be, therefore, regarded as the product of expansion coefficients and data matrices such as:

$$\phi_{k\circ l}^{[i]} = \hat{Q}_{k\circ l} a_i \tag{A2.6}$$

$$\phi_{k+l}^{[i]} = \hat{Q}_{k+l} a_i \tag{A2.7}$$

Where $\hat{Q}_{k\circ l}$ and $\hat{Q}_{k+l} \in \mathbb{C}^{M\times N_{blk}}$ and $a_i = \left[a_{i1}(f_{k+l}), \dots, a_{iN_{blk}}(f_{k+l})\right]^T$ represents the i-th vector of expansion coefficients for the (k, l) frequency doublets, with $(.)^T$ denoting the transpose. To optimally represent the sea surface characteristics in terms of integral bispectral density, the set of expansion coefficients a_1 maximising the value of $b(f_k, f_l)$ in Eq. A2.3 is defined from the numerical radius of the complex product matrix B representing the bispectral density matrix:

$$B = \hat{Q}_{k\circ l}{}^H \hat{Q}_{k+l} \tag{A2.8}$$

To seek the expansion coefficients corresponding to the largest eigenvalue λ_{max}. This method allows defining an optimal approximation of the eigenvalue characterising the integral bispectral density of the wavefield for each pair of frequency components, referred to as the mode bispectrum $\lambda_1(f_k, f_l)$. That is, the integral bispectral density is best represented by the first mode of the BMD, with other modes having a minimal impact. The peak magnitude of the optimal complex engine value $|\lambda_1(f_k, f_l)|$ for the set of frequencies f_k and f_l is analogue to the peak magnitude found in the bispectrum. Therefore, the BMD defines the modal states of the wavefield in regard to interactions between frequency components, which allows the extraction of spatial structures of phase coupling and resulting triadic interactions in two dimensions.