What can radars tell us about snowfall microphysics? Insights from a MCMC approach

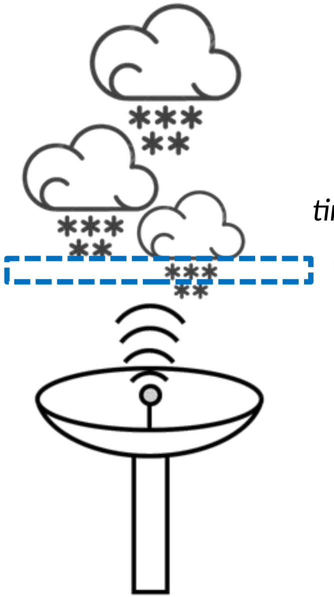
Anne-Claire Billault-Roux[1,2] and Alexis Berne[1]

1 Environmental remote sensing lab, EPFL, Lausanne, Switzerland

2Now at MeteoSwiss, Payerne, Switzerland

1st PrePEP conference, March 2025

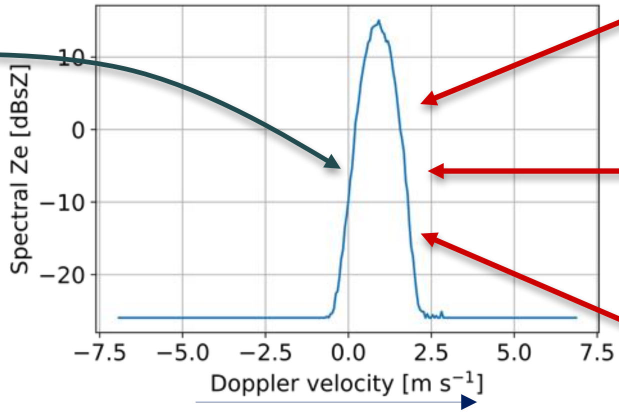
The radar Doppler spectrum



At a given time & height

Vertically-pointing radar

<u>Doppler spectrum</u>: radar reflectivity as a function of Doppler velocity



Faster-falling particles

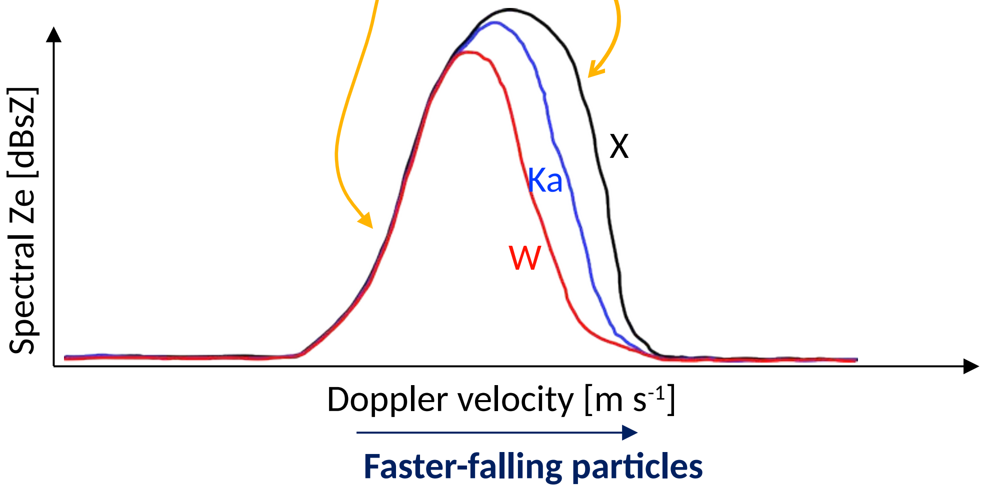
Microphysical properties
→ scattering, aerodynamics…

Radar properties
→ wavelength, beam width…

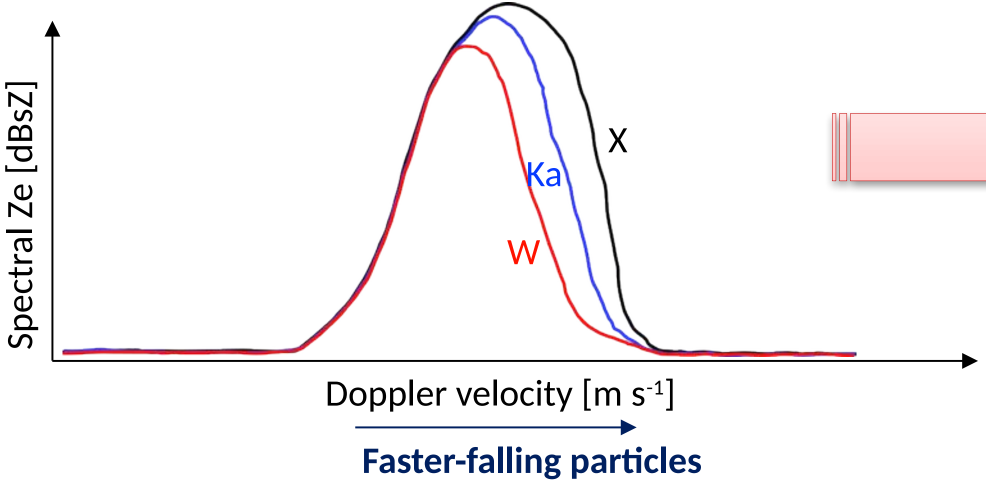
Atm. dynamic processes
→ turbulence, wind…

Multi-frequency Doppler spectra

- Rayleigh/non-Rayleigh scattering regimes: more info on PSD
  - 'Small' particles: $sZe_X \approx sZe_{Ka} \approx sZe_W$
  - 'Large' particles: $sZe_X > sZe_{Ka} > sZe_W$



Question: how much microphysical info is contained in these data?



full PSD?
PSD parameters?
density?
mass-size relation?
orientation?
geometry?
aspect ratio?
area-size relation?
...

Approach: retrieve posterior distributions using MCMC

- Given a set of triple-freq. Doppler spectra, what is the underlying distribution of microphysical descriptors that can explain the observations (**posterior distribution**)?
  - ✓ Uncertainty quantification
  - ✓ Sensitivity tests (e.g., impact of atmospheric broadening, of modelling choices…)
  - ✓ Assess reasonable assumptions for operational retrievals

Approach: retrieve posterior distributions using MCMC

- Given a set of triple-freq. Doppler spectra, what is the underlying distribution of microphysical descriptors that can explain the observations (**posterior distribution**)?
  - ✓ Uncertainty quantification
  - ✓ Sensitivity tests (e.g., impact of atmospheric broadening, of modelling choices…)
  - ✓ Assess reasonable assumptions for operational retrievals

- Markov chain Monte Carlo (**MCMC**): Bayesian retrieval framework
  - ✓ Few mathematical assumptions on retrieval problem / error distributions
  - ✓ Retrieval of full posterior distributions
  - X Too slow for use on large datasets

Framework

- MCMC, Metropolis-Hastings: iterative algorithm, based on a chain of accept/reject decisions; leading to sample full posterior distribution.

- Specific choices:
    - Adaptive proposal
    - Parallel tempering needed in some settings
    - Prior: uniform (bounds from literature)

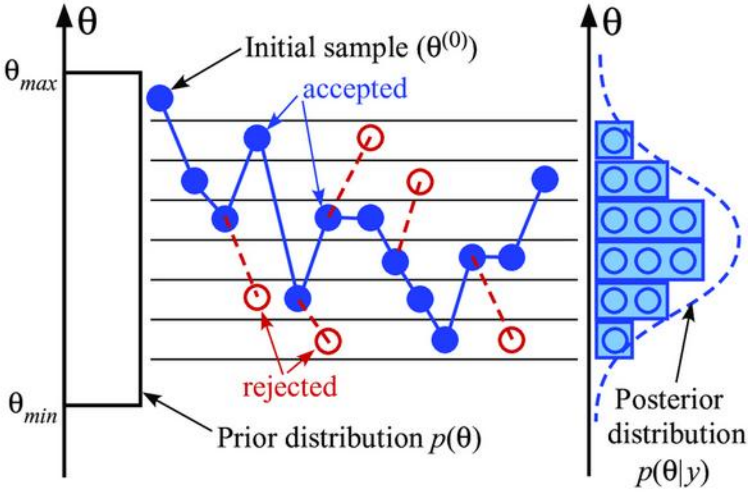


Illustration: Tomic et al. 2019

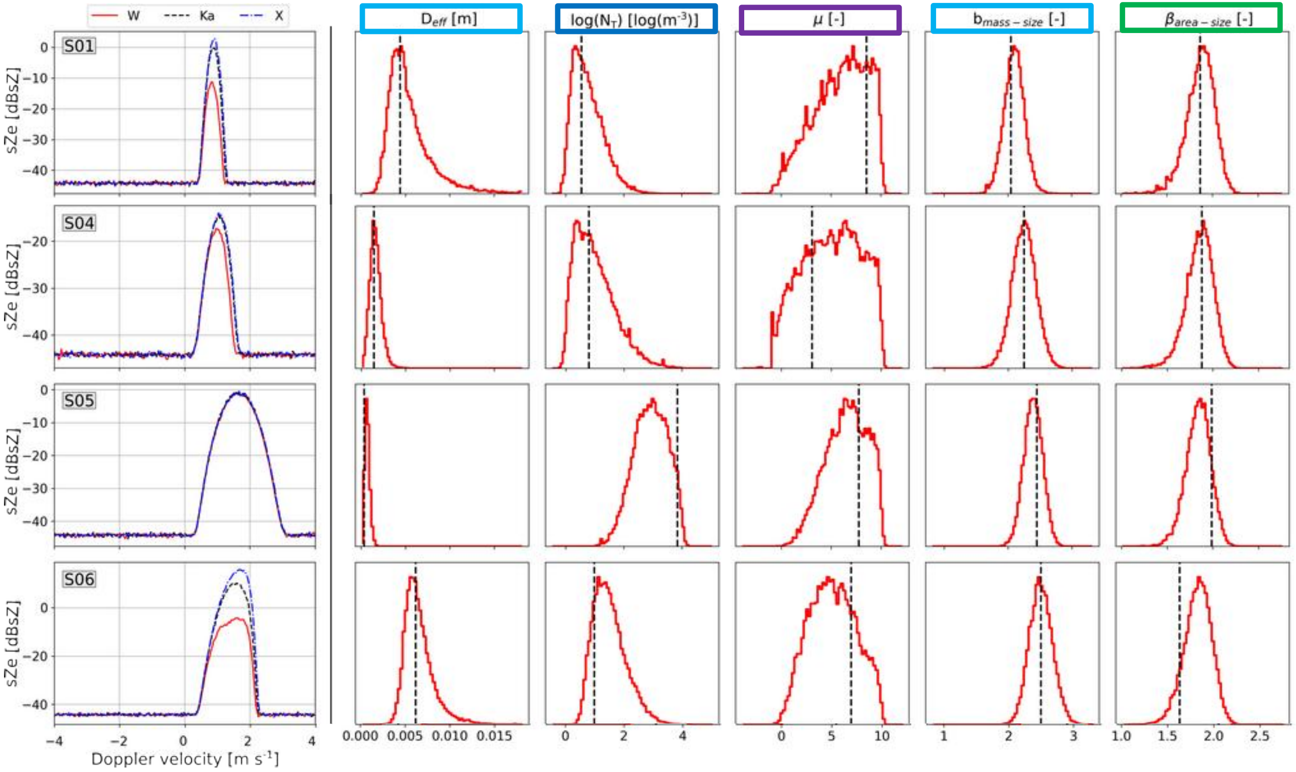
Framework

- Simulations of radar spectra using PAMTRA (Mech et al, 2020)
  - Frequencies: X + Ka + W bands
  - Scattering: Self-similar Rayleigh-Gans approximation (SSRGA, Hogan and Westbrook 2014)
  - Velocity-size: Heymsfield 2010

- Microphysical descriptors retrieved:
  - PSD shape: Gamma (prescribed through D_{eff}, μ, N_T)
  - Mass-size relation: a_m, b_m (correlated in prior)
  - Area-size relation: α_a, β_a (correlated in prior)
  - Turbulent broadening and radial wind shift: EDR, v_r (ON/OFF)

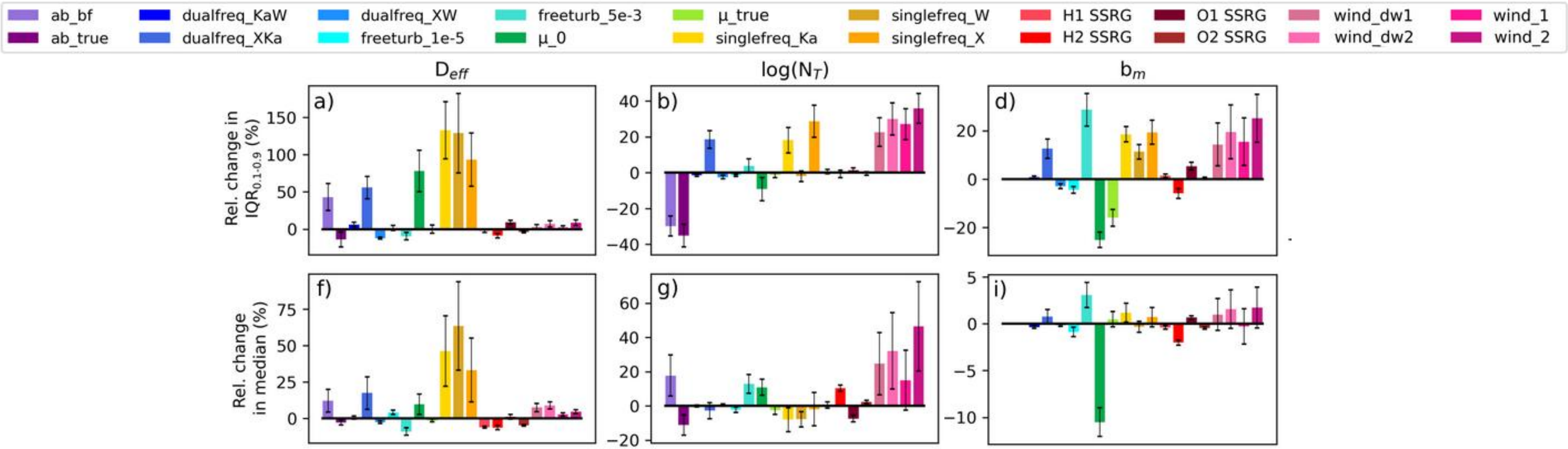
- Analysis conducted on 10 simulated examples of triple-freq. spectra (representative)

Baseline retrieval: first outcomes



- $D_{eff} + b_{mass}$:narrow + 'differentiated' posterior
- N_T : same (but log…)
- μ: wide / 'flat' posterior
- β: 'non-specific' posterior

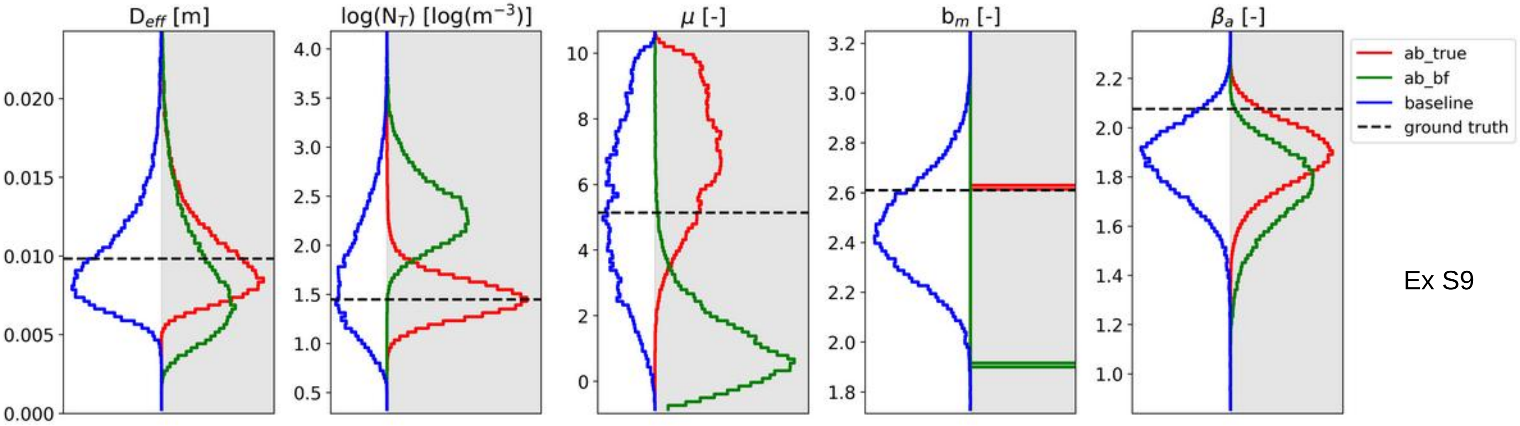
Sensitivity to retrieval parameters



- **Atmospheric effects** (turbulence + radial wind): wider and biased posterior
- Number of **frequencies**: single freq = uncertainty ++; dual-freq = worse if X+Ka only
- Using ‘wrong’ **SSRG** coefficients: minor impact
- Using ‘wrong’ **mass/size** relation: biased posterior
- Using ‘wrong’ **PSD shape** (μ): wider and biased posterior

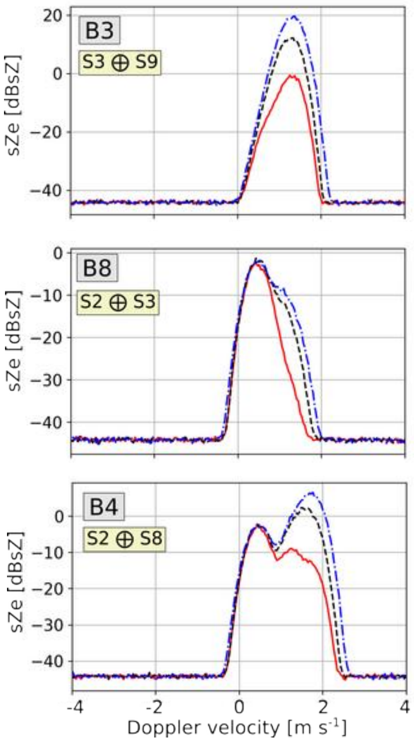
Sensitivity tests: mass-size relation

- Knowing true a_m, b_m reduces posterior width
  - e.g., - 40% $IQR_{0.1\text{-}0.9}$ in $\log(N_T)$
- Assuming wrong a_m, b_m (e.g., Brown and Francis 1995) leads to **biased** posterior
  - e.g., +20% mean in $\log(N_T)$, -100% mean in μ
  - sometimes quasi-disjoint distributions



Ex S9

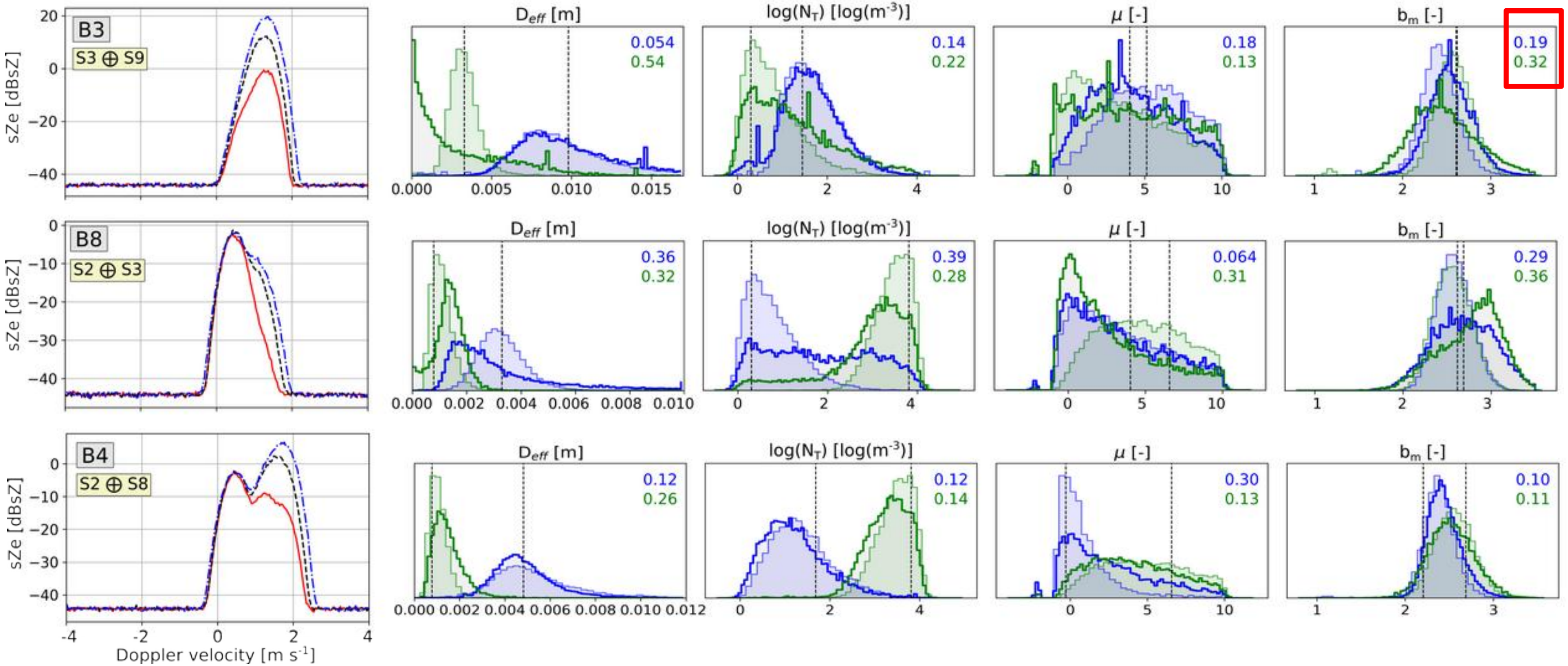
The case of bimodal Doppler spectra



- Bi-(multi-)modal spectra measured when different hydrometeor populations co-exist in radar volume
- Using MCMC framework: retrieval of 2 ice populations and compare to baseline
- Impact of bimodal “separation” on retrieval quality?

The case of bimodal Doppler spectra

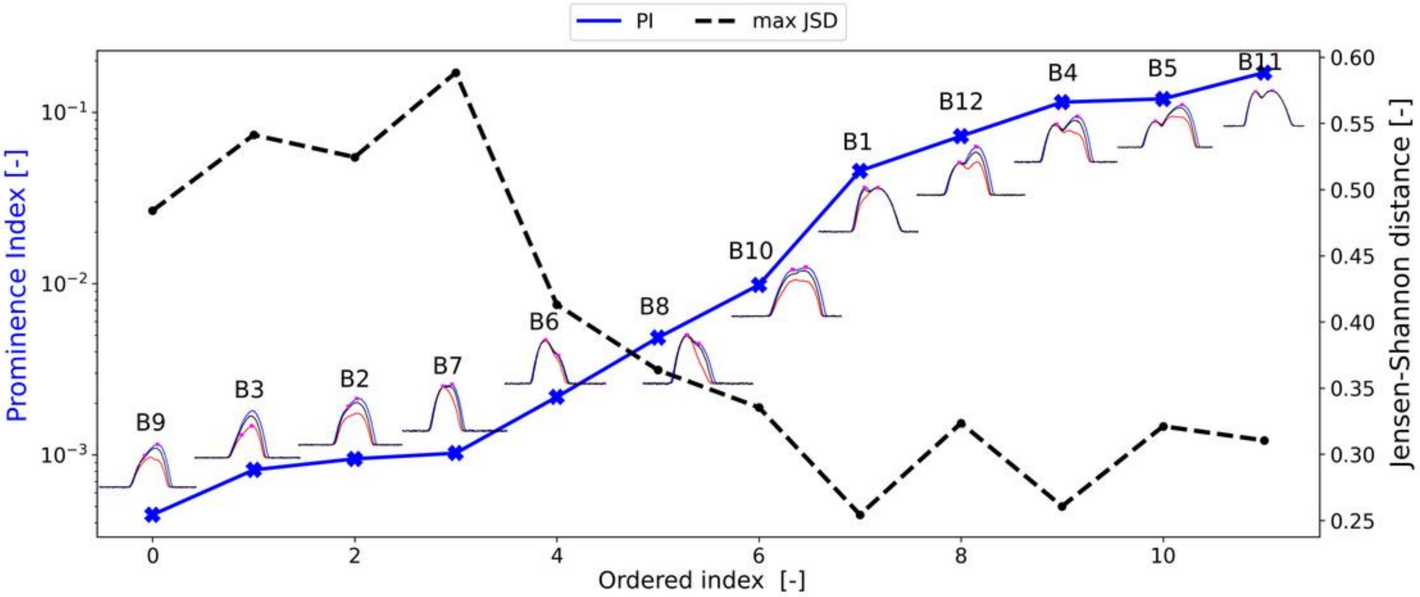
Jensen-Shannon distance w.r.t. baseline



- In some cases, distributions retrieved with similar accuracy than baseline
- Peak separation plays a great role in retrieval ability

The case of bimodal Doppler spectra

- Peak separation: quantified with prominence index (mean ratio of peak prominence) <-> related with mean JSD
- => shows information content accessible in multi-peaked spectrum



Conclusion

- MCMC framework on examples of simulated multi-frequency radar Doppler spectra
  - ✓ Few mathematical assumptions
  - ✓ Retrieval of full posterior distributions
  - ξ Too slow for operational use

Conclusion

- MCMC framework on examples of simulated multi-frequency radar Doppler spectra
  - ✓ Few mathematical assumptions
  - ✓ Retrieval of full posterior distributions
  - x Too slow for operational use

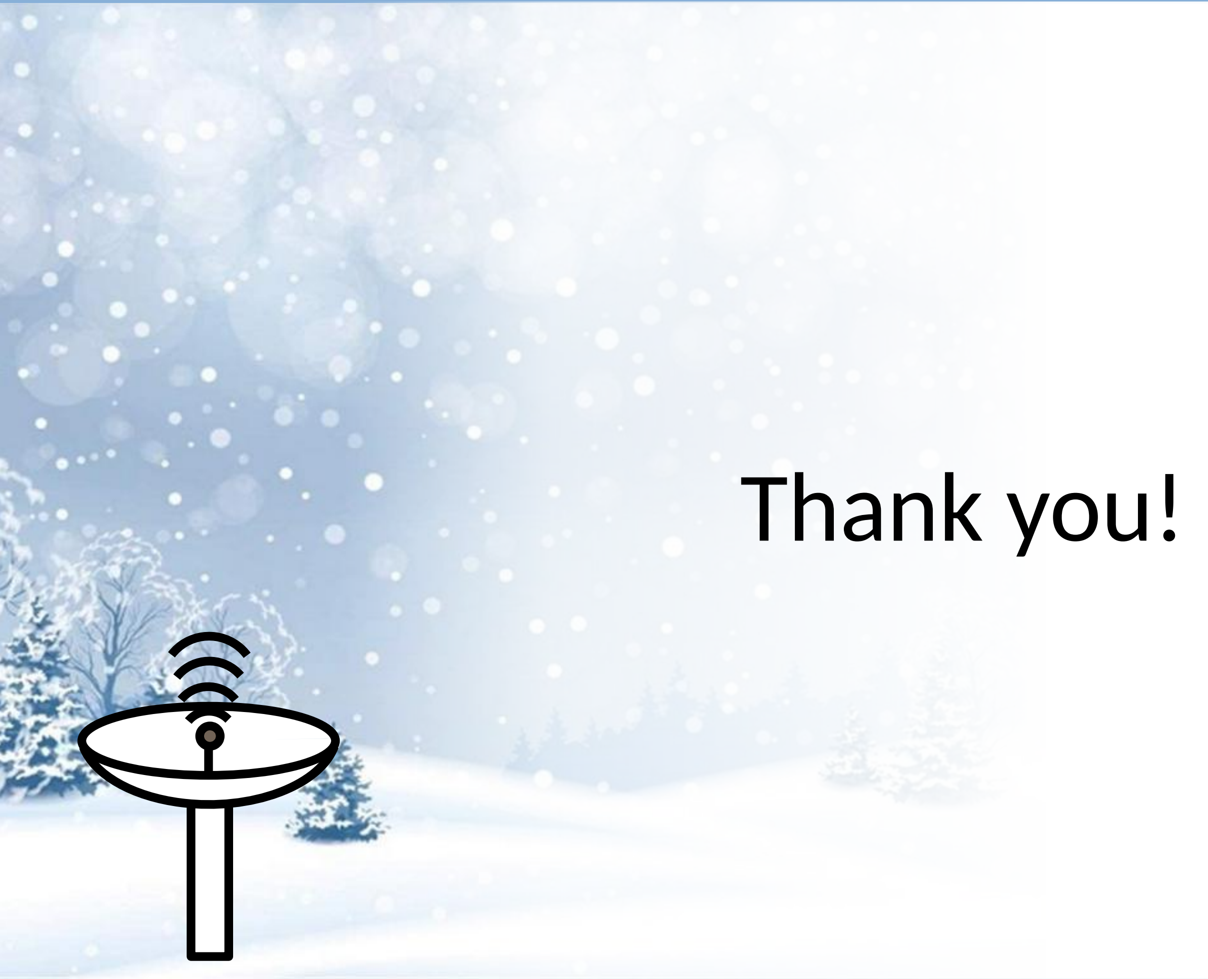
- Insights gained
  - ▸ Information content: different level of confidence depending on variable ($D_{eff} \neq \mu$)
  - ▸ Sensitivity tests: certain common assumptions can be misleading (e.g., a/b, PSD shape)
  - ▸ Bimodal spectra: prominence index indicates possibility for accurate retrieval

Conclusion

- MCMC framework on examples of simulated multi-frequency radar Doppler spectra
  - ✓ Few mathematical assumptions
  - ✓ Retrieval of full posterior distributions
  - x Too slow for operational use

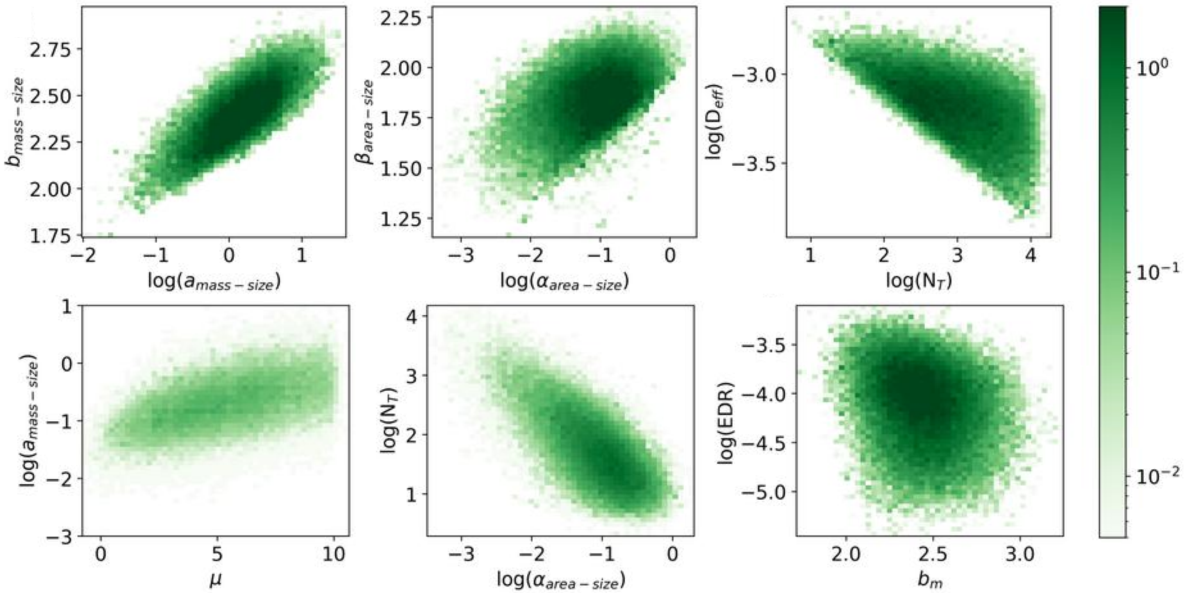
- Insights gained
  - Information content: different level of confidence depending on variable ($D_{eff} \neq \mu$)
  - Sensitivity tests: certain common assumptions can be misleading (e.g., a/b, PSD shape)
  - Bimodal spectra: prominence index indicates possibility for accurate retrieval

- Perspectives
  - Sheds light on intrinsic uncertainties
  - Using statistical emulators in view of applications on larger datasets



Thank you!

Baseline retrieval: bivariate analyses

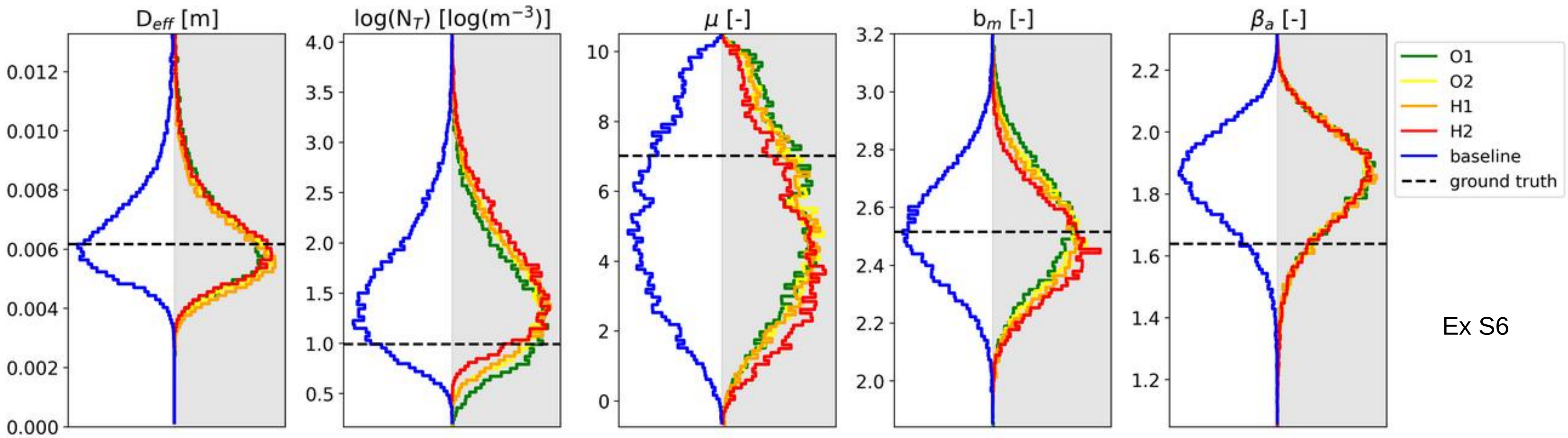


- Certain correlations come from prior ($a/b, \alpha/\beta$)
- N_T/D_{eff}: expected
- Others to be explored

Sensitivity tests: SSRG example

- Assuming a set of SSRG coefficients different than the true one (i.e., used for original simulation): only minor impact on retrieved posterior

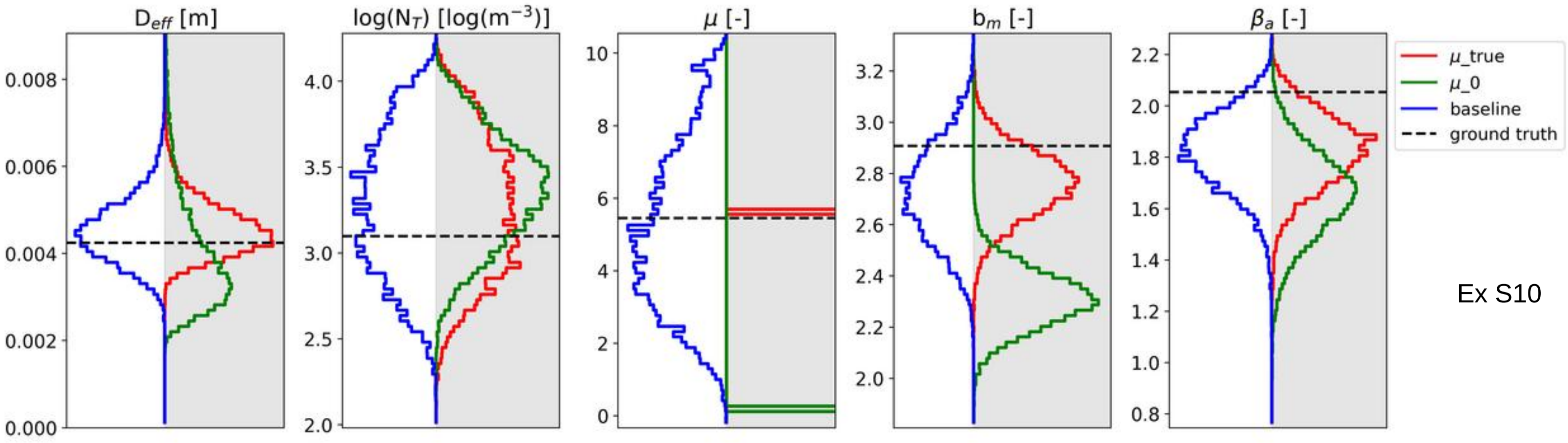
- Depending on SSRG coefs, bias in N_T and b_m



Ex S6

Sensitivity tests: PSD shape

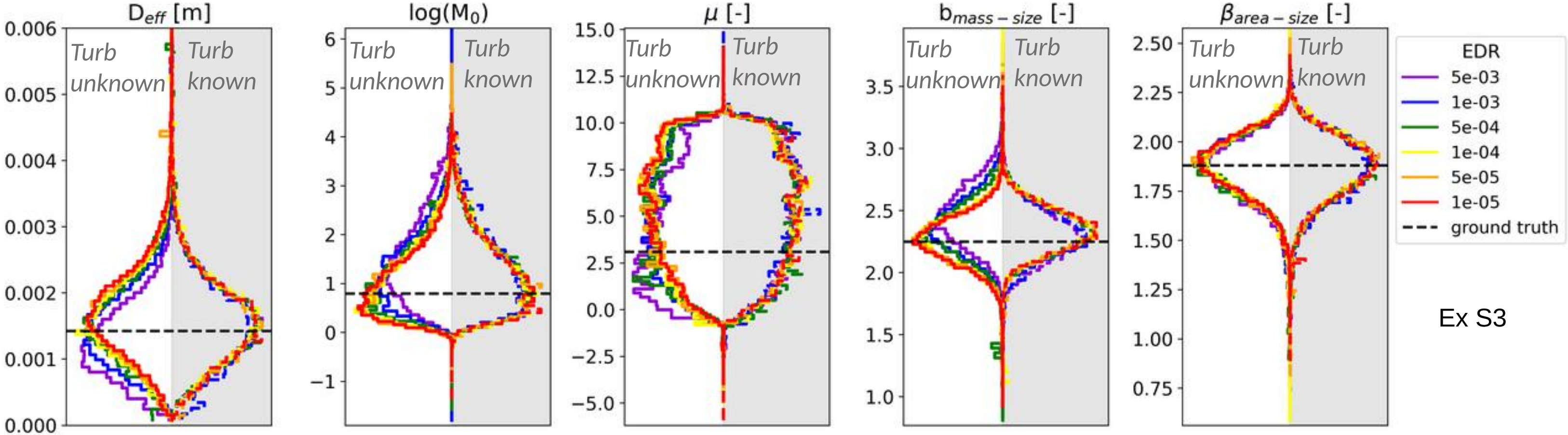
- Knowing true μ reduces posterior width (but not drastically)
- Assuming wrong μ (e.g., exp. PSD) leads to **wider** and **biased** posterior
  - e.g., +80% IQR0.1-0.9 D_{eff} , -10% mean b_m, on average
  - sometimes quasi-disjoint distributions



Ex S10

Sensitivity tests: atmospheric effects

- **Turbulent broadening**: leads to **wider** posterior and **bias**
  - e.g. +30% in $IQR_{0.1\text{-}0.9}$ for b_m on average
- But if turbulence is known: no impact on posterior (no intrinsic loss of information in spectra)

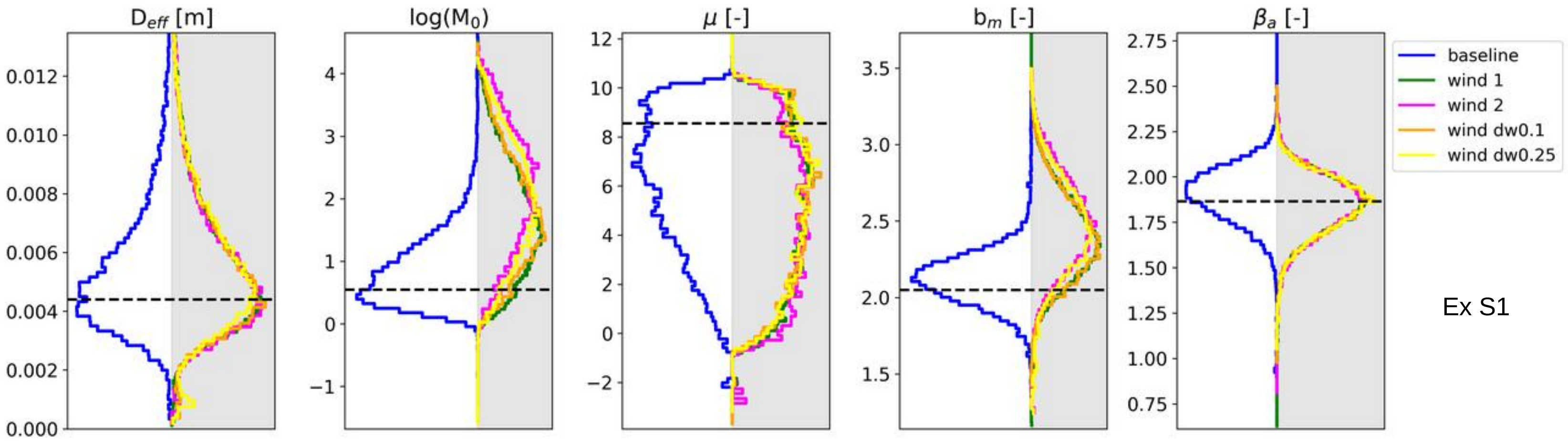


Ex S3

Sensitivity tests: atmospheric effects

- **Radial wind**: leads to **wider** posterior and **bias**
  - e.g. +40% in IQR0.1-0.9 for $\log(N_T)$ on average

- Similar effect if **uniform vertical** wind (same for 3 radars) or radar **misalignment** (differential radial wind)



Ex S1