
Hybrid Bernstein Normalizing Flows for Flexible Multivariate Density Regression with Interpretable Marginals (Supplementary Material)

Marcel Arpogaus¹

Thomas Kneib²

Thomas Nagler^{3,4}

David Rügamer^{3,4}

¹Institute for Applied Research , HTWG - University of Applied Sciences , Konstanz, GERMANY

²Chair of Statistics and Campus Institute Data Science (CIDAS) , University of Göttingen , Göttingen, GERMANY

³Department of Statistics , LMU Munich, Munich, GERMANY

⁴Munich Center for Machine Learning (MCML) , Munich, GERMANY

1 MODEL DETAILS

In the following section, we provide an overview of the most important hyper parameters used to produce our results. All models are implemented using `TensorFlow` (v2.15.1) and `TensorFlow Probability` (v0.23.0) and are trained using Adam (Dillon et al., 2017; Kingma and Ba, 2017; Martín Abadi et al., 2015). The implementation is available on GitHub¹.

1.1 SIMULATION DATA

The following tables contain all hyper parameters of the models trained on the bivariate simulation data.

¹https://github.com/MArpogaus/hybrid_flows

	CF (B)
JIT Compile	True
Epochs	400
Batch Size	512
Learning Rate Decay Steps	epochs
Initial Learning Rate	0.001
Learning Rate Scheduler	cosine_decay
Early Stopping	50
Validation Split	0.250
Output Distribution	coupling_flow
Number of Flows	2
Bijector	BernsteinPolynomial
Bernstein Bijector Extrapolation	False
Bernstein Bijector Domain Min	0
Bernstein Bijector Domain Max	1
Invert Bijector	True
Allow Flexible Bounds for Parameters	False
Bounds constraints for Parameters	linear
Nested Bernstein Bijector Codomain Min	0
Nested Bernstein Bijector Codomain Max	1
Number of Parameters	300
Low Constraint for Layer Overwrite -2	-5
High Constraint for Layer Overwrite -2	5
Low Constraint for Layer Overwrite -1	-5
High Constraint for Layer Overwrite -1	5
Activation Function	relu
Batch Normalization	False
Dropout Rate	0
First Hidden Units	128
Second Hidden Units	128
Third Hidden Units	128

Table 1: Common Hyperparameters of coupling flow (CF) (B) models trained on simulated 2D data.

Dataset Name	moons		circles	
	True	False	True	False
Conditional	True	–	True	–
Conditional Event Shape	1	–	1	–

Table 2: Specific Hyperparameters of CF (B) models trained on simulated 2D data.

	CF (S)
JIT Compile	True
Epochs	400
Batch Size	512
Learning Rate Decay Steps	epochs
Initial Learning Rate	0.001
Learning Rate Scheduler	cosine_decay
Early Stopping	50
Validation Split	0.250
Output Distribution	coupling_flow
Number of Flows	2
Bijector	RationalQuadraticSpline
Number of Parameters	95
Activation Function	relu
Batch Normalization	False
Dropout Rate	0
First Hidden Units	128
Second Hidden Units	128
Third Hidden Units	128
Spline Bijector Domain Min	-5
Interval Width for Nested Spline Bijector	10
Minimum Slope for Nested Spline Bijector	0.001
Minimum Bin Width for Nested Spline Bijector	0.001
Number of Bins for Nested Spline Bijector	32

Table 3: Common Hyperparameters of CF (S) models trained on simulated 2D data.

Dataset Name	moons		circles	
	True	False	True	False
Conditional	True	–	True	–
Conditional Event Shape	1	–	1	–

Table 4: Specific Hyperparameters of CF (S) models trained on simulated 2D data.

	HCF (B)
Two-Stage Training	True
First Marginal Bijector	BernsteinPolynomial
Invert First Marginal Bijector	True
Marginal Bijector Domain Min	0
Marginal Bijector Domain Max	1
Marginal Bijector Extrapolation	False
First Marginal Bijector Parameter Constraint Function	mctm.activations.get_thetas_constrain_fn
Allow Flexible Bounds for First Marginal Bijector	False
Bounds constraints for First Marginal Bijector	linear
First Marginal Bijector Codomain Max	5
First Marginal Bijector Codomain Min	-5
First Marginal Bijector Parameter Shape 0	2
First Marginal Bijector Parameter Shape 1	300
First Joint Bijector	RealNVP
Parameter Shape 0 for Joint Bijector	1
Marginal Epochs	1000
Marginal Batch Size	1024
Marginal Learning Rate	0.050
Marginal Reduce LR on Plateau	False
Marginal Early Stopping	50
Marginal Monitor	val_loss
Marginal Verbosity	True
Marginal Validation Split	0.250
Joint Epochs	400
Joint Batch Size	512
Joint Learning Rate Decay Steps	epochs
Joint Initial Learning Rate	0.001
Joint Learning Rate Scheduler	cosine_decay
Joint Early Stopping	50
Joint Validation Split	0.250
Masked Dimensions	1
First Nested Bijector	BernsteinPolynomial
Invert Nested Bijector	True
Nested Bernstein Bijector Domain Min	-5
Nested Bernstein Bijector Domain Max	5
Nested Bernstein Bijector Extrapolation	False
Nested Bijector Parameter Constraint Function	mctm.activations.get_thetas_constrain_fn
Bounds constraints of Nested Bernstein Bijector	linear
Nested Bernstein Bijector Codomain Max	5
Nested Bernstein Bijector Codomain Min	-5
Activation Function for Joint Bijector	relu
Batch Normalization for Joint Bijector	False
Dropout Rate for Joint Bijector	False
First Hidden Units for Joint Bijector	512
Input Shape 0 for Joint Bijector	1

Table 5: Common Hyperparameters of hybrid CF (HCF) (B) models trained on simulated 2D data.

Dataset Name	moons		circles	
Conditional	True	False	True	False
JIT Compile	True	False	True	False
Parameters Function for First Marginal Bijector	bernstein_polynomial	parameter_vector	bernstein_polynomial	parameter_vector
Joint Bijector Conditional Event Shape	1	–	1	–
Extrapolation for First Marginal Bijector	True	–	True	–
First Marginal Bijector Conditional Event Shape	1	–	1	–
Polynomial Order for First Marginal Bijector	1	–	1	–
Second Hidden Units for Joint Bijector	32	128	512	32
Third Hidden Units for Joint Bijector	512	1024	32	1024
Parameter Shape 1 for Joint Bijector	300	2000	1500	1500
Is Conditional for Joint Bijector	True	–	True	–

Table 6: Specific Hyperparameters of HCF (B) models trained on simulated 2D data.

	HCF (S)
Two-Stage Training	True
First Marginal Bijector	BernsteinPolynomial
Invert First Marginal Bijector	True
Marginal Bijector Domain Min	0
Marginal Bijector Domain Max	1
Marginal Bijector Extrapolation	False
First Marginal Bijector Parameter Constraint Function	mctm.activations.get_thetas_constrain_fn
Allow Flexible Bounds for First Marginal Bijector	False
Bounds constraints for First Marginal Bijector	linear
First Marginal Bijector Codomain Max	5
First Marginal Bijector Codomain Min	-5
First Marginal Bijector Parameter Shape 0	2
First Marginal Bijector Parameter Shape 1	300
First Joint Bijector	RealNVP
Parameter Shape 0 for Joint Bijector	1
Marginal Epochs	1000
Marginal Batch Size	1024
Marginal Learning Rate	0.050
Marginal Reduce LR on Plateau	False
Marginal Early Stopping	50
Marginal Monitor	val_loss
Marginal Verbosity	True
Marginal Validation Split	0.250
Joint Epochs	400
Joint Batch Size	512
Joint Learning Rate Decay Steps	epochs
Joint Initial Learning Rate	0.001
Joint Learning Rate Scheduler	cosine_decay
Joint Early Stopping	50
Joint Validation Split	0.250
Masked Dimensions	1
First Nested Bijector	RationalQuadraticSpline
Activation Function for Joint Bijector	relu
Batch Normalization for Joint Bijector	False
Dropout Rate for Joint Bijector	False
First Hidden Units for Joint Bijector	128
Second Hidden Units for Joint Bijector	128
Third Hidden Units for Joint Bijector	128
Input Shape 0 for Joint Bijector	1
Parameter Shape 1 for Joint Bijector	95
Nested Spline Bijector Range Min	-5
Interval Width for Nested Spline Bijector	10
Minimum Slope for Nested Spline Bijector	0.001
Minimum Bin Width for Nested Spline Bijector	0.001
Number of Bins for Nested Spline Bijector	32

Table 7: Common Hyperparameters of HCF (S) models trained on simulated 2D data.

Dataset Name	moons		circles	
Conditional	True	False	True	False
JIT Compile	True	False	True	False
Parameters Function for First Marginal Bijector	bernstein_polynomial	parameter_vector	bernstein_polynomial	parameter_vector
Joint Bijector Conditional Event Shape	1	–	1	–
Extrapolation for First Marginal Bijector	True	–	True	–
First Marginal Bijector Conditional Event Shape	1	–	1	–
Polynomial Order for First Marginal Bijector	1	–	1	–
Is Conditional for Joint Bijector	True	–	True	–

Table 8: Specific Hyperparameters of HCF (S) models trained on simulated 2D data.

	MAF (B)
JIT Compile	False
Epochs	400
Batch Size	512
Learning Rate Decay Steps	epochs
Initial Learning Rate	0.001
Learning Rate Scheduler	cosine_decay
Early Stopping	50
Validation Split	0.250
Output Distribution	masked_autoregressive_flow
Number of Flows	2
Bijector	BernsteinPolynomial
Bernstein Bijector Extrapolation	False
Bernstein Bijector Domain Min	-5
Bernstein Bijector Domain Max	5
Invert Bijector	True
Allow Flexible Bounds for Parameters	False
Bounds constraints for Parameters	linear
Nested Bernstein Bijector Codomain Min	-5
Nested Bernstein Bijector Codomain Max	5
Number of Parameters	128
Activation Function	relu
First Hidden Units	128
Second Hidden Units	128
Third Hidden Units	128
Layer Overwrite 0 Domain Min	0
Layer Overwrite 0 Domain Max	1

Table 9: Common Hyperparameters of masked autoregressive flow (MAF) (B) models trained on simulated 2D data.

Dataset Name	moons		circles	
Conditional	True	False	True	False
Conditional	True	–	True	–
Conditional Event Shape	1	–	1	–

Table 10: Specific Hyperparameters of MAF (B) models trained on simulated 2D data.

	MAF (S)
JIT Compile	False
Epochs	400
Batch Size	512
Learning Rate Decay Steps	epochs
Initial Learning Rate	0.001
Learning Rate Scheduler	cosine_decay
Early Stopping	50
Validation Split	0.250
Output Distribution	masked_autoregressive_flow
Number of Flows	2
Bijector	RationalQuadraticSpline
Number of Parameters	95
Activation Function	relu
First Hidden Units	128
Second Hidden Units	128
Third Hidden Units	128
Spline Bijector Domain Min	-5
Interval Width for Nested Spline Bijector	10
Minimum Slope for Nested Spline Bijector	0.001
Minimum Bin Width for Nested Spline Bijector	0.001
Number of Bins for Nested Spline Bijector	32

Table 11: Common Hyperparameters of MAF (S) models trained on simulated 2D data.

Dataset Name	moons		circles	
	True	False	True	False
Conditional	True	–	True	–
Conditional Event Shape	1	–	1	–

Table 12: Specific Hyperparameters of MAF (S) models trained on simulated 2D data.

	MCTM
Epochs	100
Batch Size	512
Learning Rate Decay Steps	100
Initial Learning Rate	0.010
Learning Rate Scheduler	cosine_decay
Validation Split	0.250
Two-Stage Training	True
First Marginal Bijector	BernsteinPolynomial
Invert First Marginal Bijector	True
Marginal Bijector Domain Min	0
Marginal Bijector Domain Max	1
Marginal Bijector Extrapolation	True
First Marginal Bijector Parameter Constraint Function	mctm.activations.get_thetas_constrain_fn
Allow Flexible Bounds for First Marginal Bijector	False
Bounds constraints for First Marginal Bijector	smooth
First Marginal Bijector Codomain Max	4
First Marginal Bijector Codomain Min	-4
Parameters Function for First Marginal Bijector	parameter_vector
First Marginal Bijector Parameter Shape 0	2
First Marginal Bijector Parameter Shape 1	300
First Joint Bijector	ScaleMatvecLinearOperator
Invert First Joint Bijector	True
Parameter Shape 0 for Joint Bijector	1

Table 13: Common Hyperparameters of multivariate conditional transformation model (MCTM) models trained on simulated 2D data.

Dataset Name	moons		circles	
Conditional	True	False	True	False
Early Stopping	False	False	30	False
Second Marginal Bijector	Shift	–	Shift	–
Invert Second Marginal Bijector	True	–	True	–
Parameters Function for Second Marginal Bijector	bernstein_polynomial	–	bernstein_polynomial	–
Second Marginal Bijector Conditional Event Shape	1	–	1	–
Extrapolation for Second Marginal Bijector	True	–	True	–
Second Marginal Bijector Parameter Shape 0	2	–	2	–
Polynomial Order for Second Marginal Bijector	6	–	6	–
Joint Bijector Conditional Event Shape	1	–	1	–
Extrapolation for Joint Bijector	True	–	True	–
Polynomial Order for Joint Bijector	3	–	3	–
Parameter Shape 1 for Joint Bijector	–	1	–	1

Table 14: Specific Hyperparameters of MCTM models trained on simulated 2D data.

	MVN
JIT Compile	False
Epochs	200
Batch Size	256
Learning Rate Decay Steps	200
Initial Learning Rate	0.010
Learning Rate Scheduler	cosine_decay
Early Stopping	False
Validation Split	0.100
Output Distribution	multivariate_normal

Table 15: Common Hyperparameters of multivariate normal (MVN) models trained on simulated 2D data.

Dataset Name	moons		circles	
	True	False	True	False
Activation Function	relu	–	relu	–
Batch Normalization	True	–	True	–
Dropout Rate	0	–	0	–
First Hidden Units	16	–	16	–
Second Hidden Units	16	–	16	–
Conditional	True	False	True	False
Conditional Event Shape	1	–	1	–

Table 16: Specific Hyperparameters of MVN models trained on simulated 2D data.

1.2 BENCHMARK DATA

The following tables contain all hyper parameters of the models trained on the multivariate benchmark data.

	HMAF (S)
Two-Stage Training	True
First Marginal Bijector	BernsteinPolynomial
Invert First Marginal Bijector	True
Marginal Bijector Extrapolation	True
First Marginal Bijector Parameter Constraint Function	mctm.activations.get_thetas_constrain_fn
Allow Flexible Bounds for First Marginal Bijector	False
Bounds constraints for First Marginal Bijector	smooth
First Marginal Bijector Codomain Max	4
First Marginal Bijector Codomain Min	-4
Parameters Function for First Marginal Bijector	parameter_vector
First Marginal Bijector Parameter Shape 1	2048
Marginal Reduce LR on Plateau	False
Marginal Early Stopping	30
Marginal Monitor	val_loss
Marginal Verbosity	True
Joint Epochs	200
Joint Learning Rate Decay Steps	200
Joint Learning Rate Scheduler	cosine_decay
Joint Early Stopping	30
Type of Joint Flow	masked_autoregressive_flow_first_dim_masked
Number of Flows	10
Use Invertible Linear Transformations	True
Random Permutation Seed	1
Joint Bijector	RationalQuadraticSpline
Joint Bijector Range Min	-4
Interval Width for Joint Bijector	8
Minimum Slope for Joint Bijector	0.001
Minimum Bin Width for Joint Bijector	0.001
Number of Bins for Joint Bijector	8
Number of Parameters for Joint Bijectors	23
Activation Function for MAF	relu
Activation Function for X0 Parameters	relu
Batch Normalization for X0 Parameters	False
Dropout Rate for X0 Parameters	False

Table 17: Common Hyperparameters of hybrid MAF (HMAF) (S) models trained on benchmark data.

Dataset Name	bsds300	gas	hepmass	miniboone	power
Conditional	False	False	False	False	False
First Marginal Bijector Parameter Shape 0	63	8	21	43	6
Marginal Epochs	1000	1000	2000	2000	1000
Marginal Batch Size	16384	16384	16384	512	16384
Marginal Learning Rate	0.100	0.100	0.100	0.010	0.100
Joint Batch Size	4096	4096	4096	512	4096
Joint Initial Learning Rate	0.001	0.001	0.001	0.001	0.001
First Hidden Units for MAF	512	512	64	64	256
Second Hidden Units for MAF	512	512	64	64	256
First Hidden Units for X0 Network	512	512	64	64	256
Second Hidden Units for X0 Network	512	512	64	64	256

Table 18: Specific Hyperparameters of HMAF (S) models trained on benchmark data.

	MAF (S)
Epochs	200
Learning Rate Decay Steps	200
Learning Rate Scheduler	cosine_decay
Early Stopping	30
Output Distribution	masked_autoregressive_flow
Number of Flows	10
Bijector	RationalQuadraticSpline
Activation Function	relu
Spline Bijector Domain Min	-3
Interval Width for Nested Spline Bijector	6
Minimum Slope for Nested Spline Bijector	0.001
Minimum Bin Width for Nested Spline Bijector	0.001
Use Invertible Linear Transformations	True

Table 19: Common Hyperparameters of MAF (S) models trained on benchmark data.

Dataset Name	bsds300	gas	hepmass	miniboone	power
Conditional	False	False	False	False	False
Batch Size	4096	4096	4096	512	4096
Initial Learning Rate	0.001	0.001	0.001	0.000	0.001
Number of Parameters	23	23	23	11	11
First Hidden Units	512	256	64	64	512
Second Hidden Units	512	256	64	64	512
Number of Bins for Nested Spline Bijector	8	8	8	4	4

Table 20: Specific Hyperparameters of MAF (S) models trained on benchmark data.

2 MALNUTRITION DATA

The following tables contain all hyper parameters of the models trained on the trivariate malnutrition data.

	malnutrition
Two-Stage Training	True
First Marginal Bijector	BernsteinPolynomial
Invert First Marginal Bijector	True
Marginal Bijector Domain Min	-5
Marginal Bijector Domain Max	5
Marginal Bijector Extrapolation	True
First Marginal Bijector Parameter Constraint Function	mctm.activations.get_thetas_constrain_fn
Allow Flexible Bounds for First Marginal Bijector	False
Bounds constraints for First Marginal Bijector	linear
First Marginal Bijector Codomain Max	5
First Marginal Bijector Codomain Min	-5
Parameters Function for First Marginal Bijector	parameter_vector
First Marginal Bijector Parameter Shape 0	3
First Marginal Bijector Parameter Shape 1	6
Second Marginal Bijector	Shift
Invert Second Marginal Bijector	True
Parameters Function for Second Marginal Bijector	bernstein_polynomial
Second Marginal Bijector Conditional Event Shape	1
Extrapolation for Second Marginal Bijector	True
Second Marginal Bijector Parameter Shape 0	3
Polynomial Order for Second Marginal Bijector	6
Marginal Epochs	1000
Marginal Batch Size	1024
Marginal Learning Rate	0.050
Marginal Reduce LR on Plateau	False
Marginal Early Stopping	50
Marginal Monitor	val_loss
Marginal Verbosity	True
Joint Epochs	400
Joint Batch Size	512
Joint Learning Rate Decay Steps	epochs
Joint Learning Rate Scheduler	cosine_decay
Joint Early Stopping	50
Second Marginal Bijector Domain Min	0
Second Marginal Bijector Domain Max	35

Table 21: Common Hyperparameters of models trained on malnutrition data.

Model	MCTM	HMAF (B)	HMAF (S)
JIT Compile	–	True	True
First Joint Bijector	ScaleMatvecLinearOperator	–	–
Invert First Joint Bijector	True	–	–
Joint Bijector Conditional Event Shape	1	–	–
Extrapolation for Joint Bijector	True	–	–
Parameter Shape 0 for Joint Bijector	3	–	–
Polynomial Order for Joint Bijector	6	–	–
Joint Initial Learning Rate	0.001	0.001	0.001
Type of Joint Flow	–	masked_autoregressive_flow_first_dim_masked	masked_autoregressive_flow_first_dim_masked
Number of Flows	–	1	1
Joint Bijector	–	BernsteinPolynomial	RationalQuadraticSpline
Joint Bijector Range Min	–	–	-5
Interval Width for Joint Bijector	–	–	10
Minimum Slope for Joint Bijector	–	–	0.001
Minimum Bin Width for Joint Bijector	–	–	0.001
Number of Bins for Joint Bijector	–	–	16
Number of Parameters for Joint Bijectors	–	32	47
Activation Function for MAF	–	relu	relu
First Hidden Units for MAF	–	512	512
Second Hidden Units for MAF	–	512	512
Activation Function for X0 Parameters	–	relu	relu
Batch Normalization for X0 Parameters	–	False	False
Dropout Rate for X0 Parameters	–	False	False
First Hidden Units for X0 Network	–	512	512
Second Hidden Units for X0 Network	–	512	512
Nested Bernstein Bijector Domain Min	0	–	–
Nested Bernstein Bijector Domain Max	35	–	–
Joint Bijector Extrapolation	–	False	–
Nested Bernstein Bijector Domain Min	–	-5	–
Nested Bernstein Bijector Domain Max	–	5	–
Invert Joint Bijector	–	True	–
Joint Bijector Codomain Min	–	-5	–
Joint Bijector Codomain Max	–	5	–
Third Hidden Units for MAF	–	512	512
Is Conditional for MAF	–	True	True
MAF Conditional Event Shape	–	1	1
Third Hidden Units for X0 Network	–	512	512
Is Conditional for X0 Parameters	–	True	True
X0 Parameters Conditional Event Shape	–	1	1

Table 22: Specific Hyperparameters of models trained on malnutrition data.

3 DESCRIPTION OF HYPERPARAMETERS

The following table gives a brief description of the hyperparameters and maps their table names to the canonical path inside the configurations files in the git repository².

Table Name	Parameter / Description
JIT Compile	compile_kwargs.jit_compile Defines whether to compile the model using Just-In-Time compilation for performance
Epochs	fit_kwargs.epochs Maximum number of epochs
Batch Size	fit_kwargs.batch_size Batch size used for training
Learning Rate Decay Steps	fit_kwargs.learning_rate.scheduler_kwargs.decay_steps Steps after which to decay the learning rate with the scheduler
Initial Learning Rate	fit_kwargs.learning_rate.scheduler_kwargs.initial_learning_rate Initial learning rate for the learning rate scheduler
Learning Rate Scheduler	fit_kwargs.learning_rate.scheduler_name Name of the learning rate scheduler
Early Stopping	fit_kwargs.early_stopping Enable early stopping
Validation Split	fit_kwargs.validation_split Fraction of training data to be used as validation data
Output Distribution	model_kwargs.distribution Specifies the type of output distribution used in the model
Number of Flows	model_kwargs.num_layers Total number of layers in the model
Bijector	model_kwargs.bijector Specifies the bijector type to be used in the model
Bernstein Bijector Extrapolation	model_kwargs.bijector_kwargs.extrapolation Defines whether to allow extrapolation in the bijector
Bernstein Bijector Domain Min	model_kwargs.bijector_kwargs.domain.0 Lower bound of the Bernstein polynomial bijector domain
Bernstein Bijector Domain Max	model_kwargs.bijector_kwargs.domain.1 Upper bound of the Bernstein polynomial bijector domain
Invert Bijector	model_kwargs.invert Indicates if the bijector should be inverted
Allow Flexible Bounds for Parameters	model_kwargs.parameters_constraint_fn_kwargs.allow_flexible_bounds

Continued on next page

²https://github.com/MArpogaus/hybrid_flows

Table Name	Parameter / Description
	Defines if flexible bounds are allowed for the parameters
Bounds constraints for Parameters	model_kwargs.parameters_constraint_fn_kwargs.bounds Defining the type of constraints on the bounds for the parameters
Nested Bernstein Bijector Codomain Min	model_kwargs.parameters_constraint_fn_kwargs.low Defines the lower bound for the corresponding parameters
Nested Bernstein Bijector Codomain Max	model_kwargs.parameters_constraint_fn_kwargs.high Defines the upper bound for the corresponding parameters
Number of Parameters	model_kwargs.num_parameters Total number of parameters in the model
Low Constraint for Layer Overwrite -2	model_kwargs.layer_overwrites.-2.parameters_constraint_fn_kwargs.low Lower constraint for parameters in layer overwrite -2
High Constraint for Layer Overwrite -2	model_kwargs.layer_overwrites.-2.parameters_constraint_fn_kwargs.high Upper constraint for parameters in layer overwrite -2
Low Constraint for Layer Overwrite -1	model_kwargs.layer_overwrites.-1.parameters_constraint_fn_kwargs.low Lower constraint for parameters in layer overwrite -1
High Constraint for Layer Overwrite -1	model_kwargs.layer_overwrites.-1.parameters_constraint_fn_kwargs.high Upper constraint for parameters in layer overwrite -1
Activation Function	model_kwargs.parameters_fn_kwargs.activation Activation function of the used neural network
Batch Normalization	model_kwargs.parameters_fn_kwargs.batch_norm Specifies if batch normalization should be enabled in the used neural network
Dropout Rate	model_kwargs.parameters_fn_kwargs.dropout Dropout rate for the neural network
First Hidden Units	model_kwargs.parameters_fn_kwargs.hidden_units.0 Number of units in the first hidden layer of the neural network
Second Hidden Units	model_kwargs.parameters_fn_kwargs.hidden_units.1 Number of units in the second hidden layer of the neural network
Third Hidden Units	model_kwargs.parameters_fn_kwargs.hidden_units.2 Number of units in the third hidden layer of the neural network
Conditional	model_kwargs.parameters_fn_kwargs.conditional Indicates if the parameter functions should get the covariates as input
Conditional Event Shape	model_kwargs.parameters_fn_kwargs.conditional_event_shape Shape of the covariate vector
JIT Compile	compile_kwargs.jit_compile Defines whether to compile the model using Just-In-Time compilation for performance
Epochs	fit_kwargs.epochs Maximum number of epochs

Continued on next page

Table Name	Parameter / Description
Batch Size	fit_kwargs.batch_size Batch size used for training
Learning Rate Decay Steps	fit_kwargs.learning_rate.scheduler_kwargs.decay_steps Steps after which to decay the learning rate with the scheduler
Initial Learning Rate	fit_kwargs.learning_rate.scheduler_kwargs.initial_learning_rate Initial learning rate for the learning rate scheduler
Learning Rate Scheduler	fit_kwargs.learning_rate.scheduler_name Name of the learning rate scheduler
Early Stopping	fit_kwargs.early_stopping Enable early stopping
Validation Split	fit_kwargs.validation_split Fraction of training data to be used as validation data
Output Distribution	model_kwargs.distribution Specifies the type of output distribution used in the model
Number of Flows	model_kwargs.num_layers Total number of layers in the model
Bijector	model_kwargs.bijector Specifies the bijector type to be used in the model
Number of Parameters	model_kwargs.num_parameters Total number of parameters in the model
Activation Function	model_kwargs.parameters_fn_kwargs.activation Activation function of the used neural network
Batch Normalization	model_kwargs.parameters_fn_kwargs.batch_norm Specifies if batch normalization should be enabled in the used neural network
Dropout Rate	model_kwargs.parameters_fn_kwargs.dropout Dropout rate for the neural network
First Hidden Units	model_kwargs.parameters_fn_kwargs.hidden_units.0 Number of units in the first hidden layer of the neural network
Second Hidden Units	model_kwargs.parameters_fn_kwargs.hidden_units.1 Number of units in the second hidden layer of the neural network
Third Hidden Units	model_kwargs.parameters_fn_kwargs.hidden_units.2 Number of units in the third hidden layer of the neural network
Spline Bijector Domain Min	model_kwargs.bijector_kwargs.range_min Lower bound of the quadratic spline bijector domain
Interval Width for Nested Spline Bijector	model_kwargs.parameters_constraint_fn_kwargs.interval_width Specifies the interval width for constraint parameters
Minimum Slope for Nested Spline Bijector	model_kwargs.parameters_constraint_fn_kwargs.min_slope

Continued on next page

Table Name	Parameter / Description
	Defines the minimum slope constraint for parameters
Minimum Bin Width for Nested Spline Bijector	model_kwargs.parameters_constraint_fn_kwargs.min_bin_width Defines the minimum bin width for the nested Spline bijector
Number of Bins for Nested Spline Bijector	model_kwargs.parameters_constraint_fn_kwargs.nbins Defines the number of bins used by the nested bijector
Conditional	model_kwargs.parameters_fn_kwargs.conditional Indicates if the parameter functions should get the covariates as input
Conditional Event Shape	model_kwargs.parameters_fn_kwargs.conditional_event_shape Shape of the covariate vector
JIT Compile	compile_kwargs.jit_compile Defines whether to compile the model using Just-In-Time compilation for performance
Epochs	fit_kwargs.epochs Maximum number of epochs
Batch Size	fit_kwargs.batch_size Batch size used for training
Learning Rate Decay Steps	fit_kwargs.learning_rate.scheduler_kwargs.decay_steps Steps after which to decay the learning rate with the scheduler
Initial Learning Rate	fit_kwargs.learning_rate.scheduler_kwargs.initial_learning_rate Initial learning rate for the learning rate scheduler
Learning Rate Scheduler	fit_kwargs.learning_rate.scheduler_name Name of the learning rate scheduler
Early Stopping	fit_kwargs.early_stopping Enable early stopping
Validation Split	fit_kwargs.validation_split Fraction of training data to be used as validation data
Output Distribution	model_kwargs.distribution Specifies the type of output distribution used in the model
Number of Flows	model_kwargs.num_layers Total number of layers in the model
Bijector	model_kwargs.bijector Specifies the bijector type to be used in the model
Bernstein Bijector Extrapolation	model_kwargs.bijector_kwargs.extrapolation Defines whether to allow extrapolation in the bijector
Bernstein Bijector Domain Min	model_kwargs.bijector_kwargs.domain.0 Lower bound of the Bernstein polynomial bijector domain
Bernstein Bijector Domain Max	model_kwargs.bijector_kwargs.domain.1 Upper bound of the Bernstein polynomial bijector domain

Continued on next page

Table Name	Parameter / Description
Invert Bijector	model_kwargs.invert Indicates if the bijector should be inverted
Allow Flexible Bounds for Parameters	model_kwargs.parameters_constraint_fn_kwargs.allow_flexible_bounds Defines if flexible bounds are allowed for the parameters
Bounds constraints for Parameters	model_kwargs.parameters_constraint_fn_kwargs.bounds Defining the type of constraints on the bounds for the parameters
Nested Bernstein Bijector Codomain Min	model_kwargs.parameters_constraint_fn_kwargs.low Defines the lower bound for the corresponding parameters
Nested Bernstein Bijector Codomain Max	model_kwargs.parameters_constraint_fn_kwargs.high Defines the upper bound for the corresponding parameters
Number of Parameters	model_kwargs.num_parameters Total number of parameters in the model
Activation Function	model_kwargs.parameters_fn_kwargs.activation Activation function of the used neural network
First Hidden Units	model_kwargs.parameters_fn_kwargs.hidden_units.0 Number of units in the first hidden layer of the neural network
Second Hidden Units	model_kwargs.parameters_fn_kwargs.hidden_units.1 Number of units in the second hidden layer of the neural network
Third Hidden Units	model_kwargs.parameters_fn_kwargs.hidden_units.2 Number of units in the third hidden layer of the neural network
Layer Overwrite 0 Domain Min	model_kwargs.layer_overwrites.0.bijector_kwargs.domain.0 Lower boundary for the domain in layer overwrite 0
Layer Overwrite 0 Domain Max	model_kwargs.layer_overwrites.0.bijector_kwargs.domain.1 Upper boundary for the domain in layer overwrite 0
Conditional	model_kwargs.parameters_fn_kwargs.conditional Indicates if the parameter functions should get the covariates as input
Conditional Event Shape	model_kwargs.parameters_fn_kwargs.conditional_event_shape Shape of the covariate vector
JIT Compile	compile_kwargs.jit_compile Defines whether to compile the model using Just-In-Time compilation for performance
Epochs	fit_kwargs.epochs Maximum number of epochs
Batch Size	fit_kwargs.batch_size Batch size used for training
Learning Rate Decay Steps	fit_kwargs.learning_rate.scheduler_kwargs.decay_steps Steps after which to decay the learning rate with the scheduler
Initial Learning Rate	fit_kwargs.learning_rate.scheduler_kwargs.initial_learning_rate

Continued on next page

Table Name	Parameter / Description
	Initial learning rate for the learning rate scheduler
Learning Rate Scheduler	fit_kwargs.learning_rate.scheduler_name Name of the learning rate scheduler
Early Stopping	fit_kwargs.early_stopping Enable early stopping
Validation Split	fit_kwargs.validation_split Fraction of training data to be used as validation data
Output Distribution	model_kwargs.distribution Specifies the type of output distribution used in the model
Number of Flows	model_kwargs.num_layers Total number of layers in the model
Bijector	model_kwargs.bijector Specifies the bijector type to be used in the model
Number of Parameters	model_kwargs.num_parameters Total number of parameters in the model
Activation Function	model_kwargs.parameters_fn_kwargs.activation Activation function of the used neural network
First Hidden Units	model_kwargs.parameters_fn_kwargs.hidden_units.0 Number of units in the first hidden layer of the neural network
Second Hidden Units	model_kwargs.parameters_fn_kwargs.hidden_units.1 Number of units in the second hidden layer of the neural network
Third Hidden Units	model_kwargs.parameters_fn_kwargs.hidden_units.2 Number of units in the third hidden layer of the neural network
Spline Bijector Domain Min	model_kwargs.bijector_kwargs.range_min Lower bound of the quadratic spline bijector domain
Interval Width for Nested Spline Bijector	model_kwargs.parameters_constraint_fn_kwargs.interval_width Specifies the interval width for constraint parameters
Minimum Slope for Nested Spline Bijector	model_kwargs.parameters_constraint_fn_kwargs.min_slope Defines the minimum slope constraint for parameters
Minimum Bin Width for Nested Spline Bijector	model_kwargs.parameters_constraint_fn_kwargs.min_bin_width Defines the minimum bin width for the nested Spline bijector
Number of Bins for Nested Spline Bijector	model_kwargs.parameters_constraint_fn_kwargs.nbins Defines the number of bins used by the nested bijector
Conditional	model_kwargs.parameters_fn_kwargs.conditional Indicates if the parameter functions should get the covariates as input
Conditional Event Shape	model_kwargs.parameters_fn_kwargs.conditional_event_shape Shape of the covariate vector

Continued on next page

Table Name	Parameter / Description
JIT Compile	compile_kwargs.jit_compile Defines whether to compile the model using Just-In-Time compilation for performance
Epochs	fit_kwargs.epochs Maximum number of epochs
Batch Size	fit_kwargs.batch_size Batch size used for training
Learning Rate Decay Steps	fit_kwargs.learning_rate.scheduler_kwargs.decay_steps Steps after which to decay the learning rate with the scheduler
Initial Learning Rate	fit_kwargs.learning_rate.scheduler_kwargs.initial_learning_rate Initial learning rate for the learning rate scheduler
Learning Rate Scheduler	fit_kwargs.learning_rate.scheduler_name Name of the learning rate scheduler
Early Stopping	fit_kwargs.early_stopping Enable early stopping
Validation Split	fit_kwargs.validation_split Fraction of training data to be used as validation data
Output Distribution	model_kwargs.distribution Specifies the type of output distribution used in the model
Activation Function	model_kwargs.parameters_fn_kwargs.activation Activation function of the used neural network
Batch Normalization	model_kwargs.parameters_fn_kwargs.batch_norm Specifies if batch normalization should be enabled in the used neural network
Dropout Rate	model_kwargs.parameters_fn_kwargs.dropout Dropout rate for the neural network
First Hidden Units	model_kwargs.parameters_fn_kwargs.hidden_units.0 Number of units in the first hidden layer of the neural network
Second Hidden Units	model_kwargs.parameters_fn_kwargs.hidden_units.1 Number of units in the second hidden layer of the neural network
Conditional	model_kwargs.parameters_fn_kwargs.conditional Indicates if the parameter functions should get the covariates as input
Conditional Event Shape	model_kwargs.parameters_fn_kwargs.conditional_event_shape Shape of the covariate vector
Epochs	fit_kwargs.epochs Maximum number of epochs
Batch Size	fit_kwargs.batch_size Batch size used for training
Learning Rate Decay Steps	fit_kwargs.learning_rate.scheduler_kwargs.decay_steps

Continued on next page

Table Name	Parameter / Description
	Steps after which to decay the learning rate with the scheduler
Initial Learning Rate	fit_kwargs.learning_rate.scheduler_kwargs.initial_learning_rate Initial learning rate for the learning rate scheduler
Learning Rate Scheduler	fit_kwargs.learning_rate.scheduler_name Name of the learning rate scheduler
Validation Split	fit_kwargs.validation_split Fraction of training data to be used as validation data
Two-Stage Training	two_stage_training Indicates if the model should be trained in two stages
First Marginal Bijector	model_kwargs.marginal_bijectors.0.bijector Specifies the type of the first marginal bijector
Invert First Marginal Bijector	model_kwargs.marginal_bijectors.0.invert Indicates if the first marginal bijector should allow inversion
Marginal Bijector Domain Min	model_kwargs.marginal_bijectors.0.bijector_kwargs.domain.0 Lower boundary for the domain of the first marginal bijector
Marginal Bijector Domain Max	model_kwargs.marginal_bijectors.0.bijector_kwargs.domain.1 Upper boundary for the domain of the first marginal bijector
Marginal Bijector Extrapolation	model_kwargs.marginal_bijectors.0.bijector_kwargs.extrapolation Defines whether to allow extrapolation in the first marginal bijector
First Marginal Bijector Parameter Constraint Function	model_kwargs.marginal_bijectors.0.parameters_constraint_fn Constraint function for the parameters of the first marginal bijector
Allow Flexible Bounds for First Marginal Bijector	model_kwargs.marginal_bijectors.0.parameters_constraint_fn_kwargs.allow_flexible_bounds Defines if flexible bounds are allowed for the first marginal bijector
Bounds constraints for First Marginal Bijector	model_kwargs.marginal_bijectors.0.parameters_constraint_fn_kwargs.bounds Defining the constraints on the bounds for the first marginal bijector
First Marginal Bijector Codomain Max	model_kwargs.marginal_bijectors.0.parameters_constraint_fn_kwargs.high Defines the upper bound for parameters in the first marginal bijector
First Marginal Bijector Codomain Min	model_kwargs.marginal_bijectors.0.parameters_constraint_fn_kwargs.low Defines the lower bound for parameters in the first marginal bijector
Parameters Function for First Marginal Bijector	model_kwargs.marginal_bijectors.0.parameters_fn Function for calculating parameters of the first marginal bijector
First Marginal Bijector Parameter Shape 0	model_kwargs.marginal_bijectors.0.parameters_fn_kwargs.parameter_shape.0 Defines the parameter shape dimension 0 for the first marginal bijector
First Marginal Bijector Parameter Shape 1	model_kwargs.marginal_bijectors.0.parameters_fn_kwargs.parameter_shape.1 Defines the parameter shape dimension 1 for the first marginal bijector
First Joint Bijector	model_kwargs.joint_bijectors.0.bijector Specifies the first bijector type for the joint distribution

Continued on next page

Table Name	Parameter / Description
Invert First Joint Bijector	model_kwargs.joint_bijectors.0.invert Indicates if the first joint bijector should allow inversion
Parameter Shape 0 for Joint Bijector	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.parameter_shape.0 Defines the parameter shape dimension 0 for the joint bijector
Early Stopping	fit_kwargs.early_stopping Enable early stopping
Second Marginal Bijector	model_kwargs.marginal_bijectors.1.bijector Specifies the type of second marginal bijector
Invert Second Marginal Bijector	model_kwargs.marginal_bijectors.1.invert Indicates if the second marginal bijector should allow inversion
Parameters Function for Second Marginal Bijector	model_kwargs.marginal_bijectors.1.parameters_fn Function for calculating parameters of the second marginal bijector
Second Marginal Bijector Conditional Event Shape	model_kwargs.marginal_bijectors.1.parameters_fn_kwargs.conditional_event_shape Shape of the conditional events for the second marginal bijector
Extrapolation for Second Marginal Bijector	model_kwargs.marginal_bijectors.1.parameters_fn_kwargs.extrapolation Defines whether to allow extrapolation in the second marginal bijector
Second Marginal Bijector Parameter Shape 0	model_kwargs.marginal_bijectors.1.parameters_fn_kwargs.parameter_shape.0 Defines the parameter shape dimension 0 for the second marginal bijector
Polynomial Order for Second Marginal Bijector	model_kwargs.marginal_bijectors.1.parameters_fn_kwargs.polynomial_order Defines the order of the polynomial for computations in the second marginal bijector
Joint Bijector Conditional Event Shape	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.conditional_event_shape Shape of the conditional events for the joint bijector
Extrapolation for Joint Bijector	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.extrapolation Defines whether to allow extrapolation in the joint bijector
Polynomial Order for Joint Bijector	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.polynomial_order Defines the order of the polynomial for computations in the joint bijector
Parameter Shape 1 for Joint Bijector	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.parameter_shape.1 Defines the parameter shape dimension 1 for the joint bijector
Two-Stage Training	two_stage_training Indicates if the model should be trained in two stages
First Marginal Bijector	model_kwargs.marginal_bijectors.0.bijector Specifies the type of the first marginal bijector
Invert First Marginal Bijector	model_kwargs.marginal_bijectors.0.invert Indicates if the first marginal bijector should allow inversion
Marginal Bijector Domain Min	model_kwargs.marginal_bijectors.0.bijector_kwargs.domain.0 Lower boundary for the domain of the first marginal bijector
Marginal Bijector Domain Max	model_kwargs.marginal_bijectors.0.bijector_kwargs.domain.1

Continued on next page

Table Name	Parameter / Description
	Upper boundary for the domain of the first marginal bijector
Marginal Bijector Extrapolation	model_kwargs.marginal_bijectors.0.bijector_kwargs.extrapolation Defines whether to allow extrapolation in the first marginal bijector
First Marginal Bijector Parameter Constraint Function	model_kwargs.marginal_bijectors.0.parameters_constraint_fn Constraint function for the parameters of the first marginal bijector
Allow Flexible Bounds for First Marginal Bijector	model_kwargs.marginal_bijectors.0.parameters_constraint_fn_kwargs.allow_flexible_bounds Defines if flexible bounds are allowed for the first marginal bijector
Bounds constraints for First Marginal Bijector	model_kwargs.marginal_bijectors.0.parameters_constraint_fn_kwargs.bounds Defining the constraints on the bounds for the first marginal bijector
First Marginal Bijector Codomain Max	model_kwargs.marginal_bijectors.0.parameters_constraint_fn_kwargs.high Defines the upper bound for parameters in the first marginal bijector
First Marginal Bijector Codomain Min	model_kwargs.marginal_bijectors.0.parameters_constraint_fn_kwargs.low Defines the lower bound for parameters in the first marginal bijector
First Marginal Bijector Parameter Shape 0	model_kwargs.marginal_bijectors.0.parameters_fn_kwargs.parameter_shape.0 Defines the parameter shape dimension 0 for the first marginal bijector
First Marginal Bijector Parameter Shape 1	model_kwargs.marginal_bijectors.0.parameters_fn_kwargs.parameter_shape.1 Defines the parameter shape dimension 1 for the first marginal bijector
First Joint Bijector	model_kwargs.joint_bijectors.0.bijector Specifies the first bijector type for the joint distribution
Parameter Shape 0 for Joint Bijector	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.parameter_shape.0 Defines the parameter shape dimension 0 for the joint bijector
Marginal Epochs	fit_kwargs.0.epochs Number of epochs for training the marginal distribution model
Marginal Batch Size	fit_kwargs.0.batch_size Batch size for training the marginal distribution model
Marginal Learning Rate	fit_kwargs.0.learning_rate Learning rate used to optimize the marginal distribution
Marginal Reduce LR on Plateau	fit_kwargs.0.reduce_lr_on_plateau Whether to reduce learning rate on a plateau during training
Marginal Early Stopping	fit_kwargs.0.early_stopping Enables early stopping for the marginal model training
Marginal Monitor	fit_kwargs.0.monitor Metric to monitor during training for early stopping
Marginal Verbosity	fit_kwargs.0.verbose Verbosity mode for the marginal model training
Marginal Validation Split	fit_kwargs.0.validation_split Fraction of training data to be used as validation data for the marginal model

Continued on next page

Table Name	Parameter / Description
Joint Epochs	fit_kwargs.1.epochs Maximum number of epochs for training the joint distribution model
Joint Batch Size	fit_kwargs.1.batch_size Batch size for training the joint distribution model
Joint Learning Rate Decay Steps	fit_kwargs.1.learning_rate.scheduler_kwargs.decay_steps Total number of decay steps for the learning rate scheduler of the joint model
Joint Initial Learning Rate	fit_kwargs.1.learning_rate.scheduler_kwargs.initial_learning_rate Initial learning rate for the learning rate scheduler for the joint model
Joint Learning Rate Scheduler	fit_kwargs.1.learning_rate.scheduler_name Name of the learning rate scheduler for the joint model
Joint Early Stopping	fit_kwargs.1.early_stopping Enables early stopping for the joint model training
Joint Validation Split	fit_kwargs.1.validation_split Fraction of training data to be used as validation data for the joint model
Masked Dimensions	model_kwargs.joint_bijectors.0.bijector_kwargs.num_masked Number of dimensions mask in the Coupling Flow
First Nested Bijector	model_kwargs.joint_bijectors.0.nested_bijector.bijector Specifies the type of nested bijector used in the first joint bijector
Invert Nested Bijector	model_kwargs.joint_bijectors.0.nested_bijector.invert Specifies if the nested bijector should be invertible
Nested Bernstein Bijector Domain Min	model_kwargs.joint_bijectors.0.nested_bijector.bijector_kwargs.domain.0 Lower boundary for the nested Bernstein bijector domain
Nested Bernstein Bijector Domain Max	model_kwargs.joint_bijectors.0.nested_bijector.bijector_kwargs.domain.1 Upper boundary for the nested Bernstein bijector domain
Nested Bernstein Bijector Extrapolation	model_kwargs.joint_bijectors.0.nested_bijector.bijector_kwargs.extrapolation Defines whether to allow extrapolation in the nested Bernstein bijector
Nested Bijector Parameter Constraint Function	model_kwargs.joint_bijectors.0.nested_bijector.parameters_constraint_fn Constraint function applied to the parameters of the nested bijector
Bounds constraints of Nested Bernstein Bijector	model_kwargs.joint_bijectors.0.nested_bijector.parameters_constraint_fn_kwargs.bounds Defining the type of constraints on the bounds of the nested Bernstein bijector
Nested Bernstein Bijector Codomain Max	model_kwargs.joint_bijectors.0.nested_bijector.parameters_constraint_fn_kwargs.high Defines the upper bound for the corresponding parameters
Nested Bernstein Bijector Codomain Min	model_kwargs.joint_bijectors.0.nested_bijector.parameters_constraint_fn_kwargs.low Defines the lower bound for the corresponding parameters
Activation Function for Joint Bijector	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.activation Activation function used in the neural network used to estimate the joint bijector parameters
Batch Normalization for Joint Bijector	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.batch_norm

Continued on next page

Table Name	Parameter / Description
	Specifies if batch normalization should be applied in the joint bijector
Dropout Rate for Joint Bijector	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.dropout Dropout rate to apply in the joint bijector neural network
First Hidden Units for Joint Bijector	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.hidden_units.0 Number of units in the first hidden layer of the joint bijector
Input Shape 0 for Joint Bijector	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.input_shape.0 Defines the input shape dimension 0 for the joint bijector
JIT Compile	compile_kwargs.jit_compile Defines whether to compile the model using Just-In-Time compilation for performance
Parameters Function for First Marginal Bijector	model_kwargs.marginal_bijectors.0.parameters_fn Function for calculating parameters of the first marginal bijector
Joint Bijector Conditional Event Shape	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.conditional_event_shape Shape of the conditional events for the joint bijector
Extrapolation for First Marginal Bijector	model_kwargs.marginal_bijectors.0.parameters_fn_kwargs.extrapolation Defines whether to allow extrapolation in the first marginal bijector
First Marginal Bijector Conditional Event Shape	model_kwargs.marginal_bijectors.0.parameters_fn_kwargs.conditional_event_shape Shape of conditional events for the first marginal bijector
Polynomial Order for First Marginal Bijector	model_kwargs.marginal_bijectors.0.parameters_fn_kwargs.polynomial_order Defines the order of the polynomial for computations in the first marginal bijector
Second Hidden Units for Joint Bijector	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.hidden_units.1 Number of units in the second hidden layer of the joint bijector
Third Hidden Units for Joint Bijector	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.hidden_units.2 Number of units in the third hidden layer of the joint bijector
Parameter Shape 1 for Joint Bijector	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.parameter_shape.1 Defines the parameter shape dimension 1 for the joint bijector
Is Conditional for Joint Bijector	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.conditional Indicates if the joint bijector is conditional
Two-Stage Training	two_stage_training Indicates if the model should be trained in two stages
First Marginal Bijector	model_kwargs.marginal_bijectors.0.bijector Specifies the type of the first marginal bijector
Invert First Marginal Bijector	model_kwargs.marginal_bijectors.0.invert Indicates if the first marginal bijector should allow inversion
Marginal Bijector Domain Min	model_kwargs.marginal_bijectors.0.bijector_kwargs.domain.0 Lower boundary for the domain of the first marginal bijector
Marginal Bijector Domain Max	model_kwargs.marginal_bijectors.0.bijector_kwargs.domain.1 Upper boundary for the domain of the first marginal bijector

Continued on next page

Table Name	Parameter / Description
Marginal Bijector Extrapolation	model_kwargs.marginal_bijectors.0.bijector_kwargs.extrapolation Defines whether to allow extrapolation in the first marginal bijector
First Marginal Bijector Parameter Constraint Function	model_kwargs.marginal_bijectors.0.parameters_constraint_fn Constraint function for the parameters of the first marginal bijector
Allow Flexible Bounds for First Marginal Bijector	model_kwargs.marginal_bijectors.0.parameters_constraint_fn_kwargs.allow_flexible_bounds Defines if flexible bounds are allowed for the first marginal bijector
Bounds constraints for First Marginal Bijector	model_kwargs.marginal_bijectors.0.parameters_constraint_fn_kwargs.bounds Defining the constraints on the bounds for the first marginal bijector
First Marginal Bijector Codomain Max	model_kwargs.marginal_bijectors.0.parameters_constraint_fn_kwargs.high Defines the upper bound for parameters in the first marginal bijector
First Marginal Bijector Codomain Min	model_kwargs.marginal_bijectors.0.parameters_constraint_fn_kwargs.low Defines the lower bound for parameters in the first marginal bijector
First Marginal Bijector Parameter Shape 0	model_kwargs.marginal_bijectors.0.parameters_fn_kwargs.parameter_shape.0 Defines the parameter shape dimension 0 for the first marginal bijector
First Marginal Bijector Parameter Shape 1	model_kwargs.marginal_bijectors.0.parameters_fn_kwargs.parameter_shape.1 Defines the parameter shape dimension 1 for the first marginal bijector
First Joint Bijector	model_kwargs.joint_bijectors.0.bijector Specifies the first bijector type for the joint distribution
Parameter Shape 0 for Joint Bijector	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.parameter_shape.0 Defines the parameter shape dimension 0 for the joint bijector
Marginal Epochs	fit_kwargs.0.epochs Number of epochs for training the marginal distribution model
Marginal Batch Size	fit_kwargs.0.batch_size Batch size for training the marginal distribution model
Marginal Learning Rate	fit_kwargs.0.learning_rate Learning rate used to optimize the marginal distribution
Marginal Reduce LR on Plateau	fit_kwargs.0.reduce_lr_on_plateau Whether to reduce learning rate on a plateau during training
Marginal Early Stopping	fit_kwargs.0.early_stopping Enables early stopping for the marginal model training
Marginal Monitor	fit_kwargs.0.monitor Metric to monitor during training for early stopping
Marginal Verbosity	fit_kwargs.0.verbose Verbosity mode for the marginal model training
Marginal Validation Split	fit_kwargs.0.validation_split Fraction of training data to be used as validation data for the marginal model
Joint Epochs	fit_kwargs.1.epochs

Continued on next page

Table Name	Parameter / Description
	Maximum number of epochs for training the joint distribution model
Joint Batch Size	fit_kwargs.1.batch_size Batch size for training the joint distribution model
Joint Learning Rate Decay Steps	fit_kwargs.1.learning_rate.scheduler_kwargs.decay_steps Total number of decay steps for the learning rate scheduler of the joint model
Joint Initial Learning Rate	fit_kwargs.1.learning_rate.scheduler_kwargs.initial_learning_rate Initial learning rate for the learning rate scheduler for the joint model
Joint Learning Rate Scheduler	fit_kwargs.1.learning_rate.scheduler_name Name of the learning rate scheduler for the joint model
Joint Early Stopping	fit_kwargs.1.early_stopping Enables early stopping for the joint model training
Joint Validation Split	fit_kwargs.1.validation_split Fraction of training data to be used as validation data for the joint model
Masked Dimensions	model_kwargs.joint_bijectors.0.bijector_kwargs.num_masked Number of deimensions mask in the Coupling Flow
First Nested Bijector	model_kwargs.joint_bijectors.0.nested_bijector.bijector Specifies the type of nested bijector used in the first joint bijector
Activation Function for Joint Bijector	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.activation Activation function used in the neural network used to estimate the joint bijector parameters
Batch Normalization for Joint Bijector	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.batch_norm Specifies if batch normalization should be applied in the joint bijector
Dropout Rate for Joint Bijector	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.dropout Dropout rate to apply in the joint bijector neural network
First Hidden Units for Joint Bijector	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.hidden_units.0 Number of units in the first hidden layer of the joint bijector
Second Hidden Units for Joint Bijector	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.hidden_units.1 Number of units in the second hidden layer of the joint bijector
Third Hidden Units for Joint Bijector	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.hidden_units.2 Number of units in the third hidden layer of the joint bijector
Input Shape 0 for Joint Bijector	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.input_shape.0 Defines the input shape dimension 0 for the joint bijector
Parameter Shape 1 for Joint Bijector	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.parameter_shape.1 Defines the parameter shape dimension 1 for the joint bijector
Nested Spline Bijector Range Min	model_kwargs.joint_bijectors.0.nested_bijector.bijector_kwargs.range_min Minimum range value for the nested bijector transformation
Interval Width for Nested Spline Bijector	model_kwargs.joint_bijectors.0.nested_bijector.parameters_constraint_fn_kwargs.interval_width Specifies the interval width for constraint parameters

Continued on next page

Table Name	Parameter / Description
Minimum Slope for Nested Spline Bijector	model_kwargs.joint_bijectors.0.nested_bijector.parameters_constraint_fn_kwargs.min_slope Defines the minimum slope constraint for parameters
Minimum Bin Width for Nested Spline Bijector	model_kwargs.joint_bijectors.0.nested_bijector.parameters_constraint_fn_kwargs.min_bin_width Defines the minimum bin width for the nested Spline bijector
Number of Bins for Nested Spline Bijector	model_kwargs.joint_bijectors.0.nested_bijector.parameters_constraint_fn_kwargs.nbins Defines the number of bins used by the nested bijector
JIT Compile	compile_kwargs.jit_compile Defines whether to compile the model using Just-In-Time compilation for performance
Parameters Function for First Marginal Bijector	model_kwargs.marginal_bijectors.0.parameters_fn Function for calculating parameters of the first marginal bijector
Joint Bijector Conditional Event Shape	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.conditional_event_shape Shape of the conditional events for the joint bijector
Extrapolation for First Marginal Bijector	model_kwargs.marginal_bijectors.0.parameters_fn_kwargs.extrapolation Defines whether to allow extrapolation in the first marginal bijector
First Marginal Bijector Conditional Event Shape	model_kwargs.marginal_bijectors.0.parameters_fn_kwargs.conditional_event_shape Shape of conditional events for the first marginal bijector
Polynomial Order for First Marginal Bijector	model_kwargs.marginal_bijectors.0.parameters_fn_kwargs.polynomial_order Defines the order of the polynomial for computations in the first marginal bijector
Is Conditional for Joint Bijector	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.conditional Indicates if the joint bijector is conditional
Epochs	fit_kwargs.epochs Maximum number of epochs
Learning Rate Decay Steps	fit_kwargs.learning_rate.scheduler_kwargs.decay_steps Steps after which to decay the learning rate with the scheduler
Learning Rate Scheduler	fit_kwargs.learning_rate.scheduler_name Name of the learning rate scheduler
Early Stopping	fit_kwargs.early_stopping Enable early stopping
Output Distribution	model_kwargs.distribution Specifies the type of output distribution used in the model
Number of Flows	model_kwargs.num_layers Total number of layers in the model
Bijector	model_kwargs.bijector Specifies the bijector type to be used in the model
Activation Function	model_kwargs.parameters_fn_kwargs.activation Activation function of the used neural network
Spline Bijector Domain Min	model_kwargs.bijector_kwargs.range_min

Continued on next page

Table Name	Parameter / Description
	Lower bound of the quadratic spline bijector domain
Interval Width for Nested Spline Bijector	model_kwargs.parameters_constraint_fn_kwargs.interval_width Specifies the interval width for constraint parameters
Minimum Slope for Nested Spline Bijector	model_kwargs.parameters_constraint_fn_kwargs.min_slope Defines the minimum slope constraint for parameters
Minimum Bin Width for Nested Spline Bijector	model_kwargs.parameters_constraint_fn_kwargs.min_bin_width Defines the minimum bin width for the nested Spline bijector
Use Invertible Linear Transformations	model_kwargs.use_invertible_linear_transformations Specifies if invertible linear transformations should be employed
Batch Size	fit_kwargs.batch_size Batch size used for training
Initial Learning Rate	fit_kwargs.learning_rate.scheduler_kwargs.initial_learning_rate Initial learning rate for the learning rate scheduler
Number of Parameters	model_kwargs.num_parameters Total number of parameters in the model
First Hidden Units	model_kwargs.parameters_fn_kwargs.hidden_units.0 Number of units in the first hidden layer of the neural network
Second Hidden Units	model_kwargs.parameters_fn_kwargs.hidden_units.1 Number of units in the second hidden layer of the neural network
Number of Bins for Nested Spline Bijector	model_kwargs.parameters_constraint_fn_kwargs.nbins Defines the number of bins used by the nested bijector
Two-Stage Training	two_stage_training Indicates if the model should be trained in two stages
First Marginal Bijector	model_kwargs.marginal_bijectors.0.bijector Specifies the type of the first marginal bijector
Invert First Marginal Bijector	model_kwargs.marginal_bijectors.0.invert Indicates if the first marginal bijector should allow inversion
Marginal Bijector Extrapolation	model_kwargs.marginal_bijectors.0.bijector_kwargs.extrapolation Defines whether to allow extrapolation in the first marginal bijector
First Marginal Bijector Parameter Constraint Function	model_kwargs.marginal_bijectors.0.parameters_constraint_fn Constraint function for the parameters of the first marginal bijector
Allow Flexible Bounds for First Marginal Bijector	model_kwargs.marginal_bijectors.0.parameters_constraint_fn_kwargs.allow_flexible_bounds Defines if flexible bounds are allowed for the first marginal bijector
Bounds constraints for First Marginal Bijector	model_kwargs.marginal_bijectors.0.parameters_constraint_fn_kwargs.bounds Defining the constraints on the bounds for the first marginal bijector
First Marginal Bijector Codomain Max	model_kwargs.marginal_bijectors.0.parameters_constraint_fn_kwargs.high Defines the upper bound for parameters in the first marginal bijector

Continued on next page

Table Name	Parameter / Description
First Marginal Bijector Codomain Min	model_kwargs.marginal_bijectors.0.parameters_constraint_fn_kwargs.low Defines the lower bound for parameters in the first marginal bijector
Parameters Function for First Marginal Bijector	model_kwargs.marginal_bijectors.0.parameters_fn Function for calculating parameters of the first marginal bijector
First Marginal Bijector Parameter Shape 1	model_kwargs.marginal_bijectors.0.parameters_fn_kwargs.parameter_shape.1 Defines the parameter shape dimension 1 for the first marginal bijector
Marginal Reduce LR on Plateau	fit_kwargs.0.reduce_lr_on_plateau Whether to reduce learning rate on a plateau during training
Marginal Early Stopping	fit_kwargs.0.early_stopping Enables early stopping for the marginal model training
Marginal Monitor	fit_kwargs.0.monitor Metric to monitor during training for early stopping
Marginal Verbosity	fit_kwargs.0.verbose Verbosity mode for the marginal model training
Joint Epochs	fit_kwargs.1.epochs Maximum number of epochs for training the joint distribution model
Joint Learning Rate Decay Steps	fit_kwargs.1.learning_rate.scheduler_kwargs.decay_steps Total number of decay steps for the learning rate scheduler of the joint model
Joint Learning Rate Scheduler	fit_kwargs.1.learning_rate.scheduler_name Name of the learning rate scheduler for the joint model
Joint Early Stopping	fit_kwargs.1.early_stopping Enables early stopping for the joint model training
Type of Joint Flow	model_kwargs.joint_flow_type Specifies the type of flow used in the model
Number of Flows	model_kwargs.joint_bijectors.num_layers Total number of layers in the joint bijectors
Use Invertible Linear Transformations	model_kwargs.joint_bijectors.use_invertible_linear_transformations Specifies if invertible linear transformations should be used
Random Permutation Seed	model_kwargs.joint_bijectors.random_permutation_seed Seed value for random permutation in bijectors
Joint Bijector	model_kwargs.joint_bijectors.bijector Specifies the bijector type to be used in the joint model
Joint Bijector Range Min	model_kwargs.joint_bijectors.bijector_kwargs.range_min Minimum range value for the joint bijector transformation
Interval Width for Joint Bijector	model_kwargs.joint_bijectors.parameters_constraint_fn_kwargs.interval_width Specifies the interval width for constraint parameters
Minimum Slope for Joint Bijector	model_kwargs.joint_bijectors.parameters_constraint_fn_kwargs.min_slope

Continued on next page

Table Name	Parameter / Description
	Defines the minimum slope constraint for parameters
Minimum Bin Width for Joint Bijector	model_kwargs.joint_bijectors.parameters_constraint_fn_kwargs.min_bin_width Defines the minimum bin width for the joint bijector
Number of Bins for Joint Bijector	model_kwargs.joint_bijectors.parameters_constraint_fn_kwargs.nbins Defines the number of bins used by the joint bijector
Number of Parameters for Joint Bijectors	model_kwargs.joint_bijectors.num_parameters Total number of parameters in the joint bijectors
Activation Function for MAF	model_kwargs.joint_bijectors.maf_parameters_fn_kwargs.activation Activation function used in the neural network for MAF parameters
Activation Function for X0 Parameters	model_kwargs.joint_bijectors.x0_parameters_fn_kwargs.activation Activation function for x0 parameters in the joint bijector
Batch Normalization for X0 Parameters	model_kwargs.joint_bijectors.x0_parameters_fn_kwargs.batch_norm Specifies if batch normalization should be applied to x0 parameters
Dropout Rate for X0 Parameters	model_kwargs.joint_bijectors.x0_parameters_fn_kwargs.dropout Dropout rate for the x0 parameters neural network
First Marginal Bijector Parameter Shape 0	model_kwargs.marginal_bijectors.0.parameters_fn_kwargs.parameter_shape.0 Defines the parameter shape dimension 0 for the first marginal bijector
Marginal Epochs	fit_kwargs.0.epochs Number of epochs for training the marginal distribution model
Marginal Batch Size	fit_kwargs.0.batch_size Batch size for training the marginal distribution model
Marginal Learning Rate	fit_kwargs.0.learning_rate Learning rate used to optimize the marginal distribution
Joint Batch Size	fit_kwargs.1.batch_size Batch size for training the joint distribution model
Joint Initial Learning Rate	fit_kwargs.1.learning_rate.scheduler_kwargs.initial_learning_rate Initial learning rate for the learning rate scheduler for the joint model
First Hidden Units for MAF	model_kwargs.joint_bijectors.maf_parameters_fn_kwargs.hidden_units.0 Number of units in the first hidden layer of the MAF
Second Hidden Units for MAF	model_kwargs.joint_bijectors.maf_parameters_fn_kwargs.hidden_units.1 Number of units in the second hidden layer of the MAF
First Hidden Units for X0 Network	model_kwargs.joint_bijectors.x0_parameters_fn_kwargs.hidden_units.0 Number of units in the first hidden layer for x0 parameters
Second Hidden Units for X0 Network	model_kwargs.joint_bijectors.x0_parameters_fn_kwargs.hidden_units.1 Number of units in the second hidden layer for x0 parameters
Two-Stage Training	two_stage_training Indicates if the model should be trained in two stages

Continued on next page

Table Name	Parameter / Description
First Marginal Bijector	model_kwargs.marginal_bijectors.0.bijector Specifies the type of the first marginal bijector
Invert First Marginal Bijector	model_kwargs.marginal_bijectors.0.invert Indicates if the first marginal bijector should allow inversion
Marginal Bijector Domain Min	model_kwargs.marginal_bijectors.0.bijector_kwargs.domain.0 Lower boundary for the domain of the first marginal bijector
Marginal Bijector Domain Max	model_kwargs.marginal_bijectors.0.bijector_kwargs.domain.1 Upper boundary for the domain of the first marginal bijector
Marginal Bijector Extrapolation	model_kwargs.marginal_bijectors.0.bijector_kwargs.extrapolation Defines whether to allow extrapolation in the first marginal bijector
First Marginal Bijector Parameter Constraint Function	model_kwargs.marginal_bijectors.0.parameters_constraint_fn Constraint function for the parameters of the first marginal bijector
Allow Flexible Bounds for First Marginal Bijector	model_kwargs.marginal_bijectors.0.parameters_constraint_fn_kwargs.allow_flexible_bounds Defines if flexible bounds are allowed for the first marginal bijector
Bounds constraints for First Marginal Bijector	model_kwargs.marginal_bijectors.0.parameters_constraint_fn_kwargs.bounds Defining the constraints on the bounds for the first marginal bijector
First Marginal Bijector Codomain Max	model_kwargs.marginal_bijectors.0.parameters_constraint_fn_kwargs.high Defines the upper bound for parameters in the first marginal bijector
First Marginal Bijector Codomain Min	model_kwargs.marginal_bijectors.0.parameters_constraint_fn_kwargs.low Defines the lower bound for parameters in the first marginal bijector
Parameters Function for First Marginal Bijector	model_kwargs.marginal_bijectors.0.parameters_fn Function for calculating parameters of the first marginal bijector
First Marginal Bijector Parameter Shape 0	model_kwargs.marginal_bijectors.0.parameters_fn_kwargs.parameter_shape.0 Defines the parameter shape dimension 0 for the first marginal bijector
First Marginal Bijector Parameter Shape 1	model_kwargs.marginal_bijectors.0.parameters_fn_kwargs.parameter_shape.1 Defines the parameter shape dimension 1 for the first marginal bijector
Second Marginal Bijector	model_kwargs.marginal_bijectors.1.bijector Specifies the type of second marginal bijector
Invert Second Marginal Bijector	model_kwargs.marginal_bijectors.1.invert Indicates if the second marginal bijector should allow inversion
Parameters Function for Second Marginal Bijector	model_kwargs.marginal_bijectors.1.parameters_fn Function for calculating parameters of the second marginal bijector
Second Marginal Bijector Conditional Event Shape	model_kwargs.marginal_bijectors.1.parameters_fn_kwargs.conditional_event_shape Shape of the conditional events for the second marginal bijector
Extrapolation for Second Marginal Bijector	model_kwargs.marginal_bijectors.1.parameters_fn_kwargs.extrapolation Defines whether to allow extrapolation in the second marginal bijector
Second Marginal Bijector Parameter Shape 0	model_kwargs.marginal_bijectors.1.parameters_fn_kwargs.parameter_shape.0

Continued on next page

Table Name	Parameter / Description
Polynomial Order for Second Marginal Bijector	Defines the parameter shape dimension 0 for the second marginal bijector model_kwargs.marginal_bijectors.1.parameters_fn_kwargs.polynomial_order Defines the order of the polynomial for computations in the second marginal bijector
Marginal Epochs	fit_kwargs.0.epochs Number of epochs for training the marginal distribution model
Marginal Batch Size	fit_kwargs.0.batch_size Batch size for training the marginal distribution model
Marginal Learning Rate	fit_kwargs.0.learning_rate Learning rate used to optimize the marginal distribution
Marginal Reduce LR on Plateau	fit_kwargs.0.reduce_lr_on_plateau Whether to reduce learning rate on a plateau during training
Marginal Early Stopping	fit_kwargs.0.early_stopping Enables early stopping for the marginal model training
Marginal Monitor	fit_kwargs.0.monitor Metric to monitor during training for early stopping
Marginal Verbosity	fit_kwargs.0.verbose Verbosity mode for the marginal model training
Joint Epochs	fit_kwargs.1.epochs Maximum number of epochs for training the joint distribution model
Joint Batch Size	fit_kwargs.1.batch_size Batch size for training the joint distribution model
Joint Learning Rate Decay Steps	fit_kwargs.1.learning_rate.scheduler_kwargs.decay_steps Total number of decay steps for the learning rate scheduler of the joint model
Joint Learning Rate Scheduler	fit_kwargs.1.learning_rate.scheduler_name Name of the learning rate scheduler for the joint model
Joint Early Stopping	fit_kwargs.1.early_stopping Enables early stopping for the joint model training
Second Marginal Bijector Domain Min	model_kwargs.marginal_bijectors.1.parameters_fn_kwargs.domain.0 Lower boundary for the domain of the second marginal bijector
Second Marginal Bijector Domain Max	model_kwargs.marginal_bijectors.1.parameters_fn_kwargs.domain.1 Upper boundary for the domain of the second marginal bijector
JIT Compile	compile_kwargs.jit_compile Defines whether to compile the model using Just-In-Time compilation for performance
First Joint Bijector	model_kwargs.joint_bijectors.0.bijector Specifies the first bijector type for the joint distribution
Invert First Joint Bijector	model_kwargs.joint_bijectors.0.invert Indicates if the first joint bijector should allow inversion

Continued on next page

Table Name	Parameter / Description
Joint Bijector Conditional Event Shape	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.conditional_event_shape Shape of the conditional events for the joint bijector
Extrapolation for Joint Bijector	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.extrapolation Defines whether to allow extrapolation in the joint bijector
Parameter Shape 0 for Joint Bijector	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.parameter_shape.0 Defines the parameter shape dimension 0 for the joint bijector
Polynomial Order for Joint Bijector	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.polynomial_order Defines the order of the polynomial for computations in the joint bijector
Joint Initial Learning Rate	fit_kwargs.1.learning_rate.scheduler_kwargs.initial_learning_rate Initial learning rate for the learning rate scheduler for the joint model
Type of Joint Flow	model_kwargs.joint_flow_type Specifies the type of flow used in the model
Number of Flows	model_kwargs.joint_bijectors.num_layers Total number of layers in the joint bijectors
Joint Bijector	model_kwargs.joint_bijectors.bijector Specifies the bijector type to be used in the joint model
Joint Bijector Range Min	model_kwargs.joint_bijectors.bijector_kwargs.range_min Minimum range value for the joint bijector transformation
Interval Width for Joint Bijector	model_kwargs.joint_bijectors.parameters_constraint_fn_kwargs.interval_width Specifies the interval width for constraint parameters
Minimum Slope for Joint Bijector	model_kwargs.joint_bijectors.parameters_constraint_fn_kwargs.min_slope Defines the minimum slope constraint for parameters
Minimum Bin Width for Joint Bijector	model_kwargs.joint_bijectors.parameters_constraint_fn_kwargs.min_bin_width Defines the minimum bin width for the joint bijector
Number of Bins for Joint Bijector	model_kwargs.joint_bijectors.parameters_constraint_fn_kwargs.nbins Defines the number of bins used by the joint bijector
Number of Parameters for Joint Bijectors	model_kwargs.joint_bijectors.num_parameters Total number of parameters in the joint bijectors
Activation Function for MAF	model_kwargs.joint_bijectors.maf_parameters_fn_kwargs.activation Activation function used in the neural network for MAF parameters
First Hidden Units for MAF	model_kwargs.joint_bijectors.maf_parameters_fn_kwargs.hidden_units.0 Number of units in the first hidden layer of the MAF
Second Hidden Units for MAF	model_kwargs.joint_bijectors.maf_parameters_fn_kwargs.hidden_units.1 Number of units in the second hidden layer of the MAF
Activation Function for X0 Parameters	model_kwargs.joint_bijectors.x0_parameters_fn_kwargs.activation Activation function for x0 parameters in the joint bijector
Batch Normalization for X0 Parameters	model_kwargs.joint_bijectors.x0_parameters_fn_kwargs.batch_norm

Continued on next page

Table Name	Parameter / Description
	Specifies if batch normalization should be applied to x0 parameters
Dropout Rate for X0 Parameters	model_kwargs.joint_bijectors.x0_parameters_fn_kwargs.dropout Dropout rate for the x0 parameters neural network
First Hidden Units for X0 Network	model_kwargs.joint_bijectors.x0_parameters_fn_kwargs.hidden_units.0 Number of units in the first hidden layer for x0 parameters
Second Hidden Units for X0 Network	model_kwargs.joint_bijectors.x0_parameters_fn_kwargs.hidden_units.1 Number of units in the second hidden layer for x0 parameters
Nested Bernstein Bijector Domain Min	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.domain.0 Lower boundary for the nested Bernstein bijector domain
Nested Bernstein Bijector Domain Max	model_kwargs.joint_bijectors.0.parameters_fn_kwargs.domain.1 Upper boundary for the nested Bernstein bijector domain
Joint Bijector Extrapolation	model_kwargs.joint_bijectors.bijector_kwargs.extrapolation Defines whether to allow extrapolation in the joint bijector
Nested Bernstein Bijector Domain Min	model_kwargs.joint_bijectors.bijector_kwargs.domain.0 Lower boundary for the nested Bernstein bijector domain
Nested Bernstein Bijector Domain Max	model_kwargs.joint_bijectors.bijector_kwargs.domain.1 Upper boundary for the nested Bernstein bijector domain
Invert Joint Bijector	model_kwargs.joint_bijectors.invert Indicates if the joint bijector should be invertible
Joint Bijector Codomain Min	model_kwargs.joint_bijectors.parameters_constraint_fn_kwargs.low Defines the lower bound for the corresponding parameters
Joint Bijector Codomain Max	model_kwargs.joint_bijectors.parameters_constraint_fn_kwargs.high Defines the upper bound for the corresponding parameters
Third Hidden Units for MAF	model_kwargs.joint_bijectors.maf_parameters_fn_kwargs.hidden_units.2 Number of units in the third hidden layer of the MAF
Is Conditional for MAF	model_kwargs.joint_bijectors.maf_parameters_fn_kwargs.conditional Indicates if the MAF parameters are conditional
MAF Conditional Event Shape	model_kwargs.joint_bijectors.maf_parameters_fn_kwargs.conditional_event_shape Shape of the conditional events for the MAF
Third Hidden Units for X0 Network	model_kwargs.joint_bijectors.x0_parameters_fn_kwargs.hidden_units.2 Number of units in the third hidden layer for x0 parameters
Is Conditional for X0 Parameters	model_kwargs.joint_bijectors.x0_parameters_fn_kwargs.conditional Indicates if the x0 parameters are conditional
X0 Parameters Conditional Event Shape	model_kwargs.joint_bijectors.x0_parameters_fn_kwargs.conditional_event_shape Shape of the conditional events for x0 parameters

REFERENCES

- Dillon, Joshua V. et al. (Nov. 28, 2017). *TensorFlow Distributions*. DOI: 10.48550/arXiv.1711.10604. arXiv: 1711.10604 [cs, stat]. Pre-published.
- Kingma, Diederik P. and Jimmy Ba (Jan. 29, 2017). “Adam: A Method for Stochastic Optimization.” arXiv: 1412.6980 [cs].
- Martín Abadi et al. (2015). *TensorFlow: Large-scale Machine Learning on Heterogeneous Systems*.