

スーパーカミオカンデにおける機械学習を用いた重力  
崩壊型超新星爆発の方向決定

横浜国立大学院 理工学府  
数物・電子情報系理工学専攻 物理工学  
ニュートリノ・素粒子物理学実験南野・Bronner 研究室

学籍番号 24NC240

堀口 大輔

2026年1月28日

## 概要

重力崩壊型超新星爆発 (Core-collapse Supernova Explosion: CCSNe) は、太陽質量の約 8 倍以上の質量を持つ恒星が最期に起こす爆発現象であり、その際に放出される超新星爆発ニュートリノは、可視光による超新星観測に先立って地球に到来する。このため、超新星ニュートリノバーストの観測は、早期アラートの発信を通じて、光学望遠鏡観測や重力波観測を含むマルチメッセンジャー観測における初動対応を可能にする。

スーパーカミオカンデに到来した超新星ニュートリノは、水分子との反応により生成される陽電子および電子として観測される。これらの反応のうち、大半を占める逆ベータ崩壊事象は方向再構成においてバックグラウンドとなる一方、数%含まれる電子散乱反応で生成される電子は、超新星ニュートリノの到来方向と強い相関を持つ。現行の超新星アラートシステムである SNWATCH では、ガドリニウムによる中性子タグ付けを用いた背景除去や、検出器由来バックグラウンドが多い 6MeV 以下の事象を除外することで、最尤法に基づく高精度な方向再構成を実現している。しかし、6MeV 以下の低エネルギー領域にも、方向推定に有効な電子散乱事象が多く含まれている。

そこで本研究では、事象再構成プログラム WIT を用いて低エネルギー領域も含めてニュートリノ事象の再構成を行ったデータに対して、教師なし学習の一種である密度ベースクラスタリング手法 DBSCAN を適用した新たな方向推定手法を提案する。具体的には、実際の超新星爆発時の観測事象を再現したモンテカルロシミュレーションを用い、複数の事象情報に基づくバックグラウンド除去を行った上で、本手法による方向再構成性能の評価し、早期方向決定への適用可能性を示した。

# 目次

第 1 章	重力崩壊型超新星爆発	9
1.1	爆発過程	9
1.1.1	星の進化	9
1.1.2	超新星の分類	9
1.1.3	重力崩壊	10
1.1.4	ニュートリノトラッピング	11
1.1.5	衝撃波の発生	11
1.1.6	中性子化バースト	12
1.2	マルチメッセンジャー天文学とスーパーカミオカンデ	12
第 2 章	スーパーカミオカンデにおける超新星ニュートリノバースト観測	14
2.1	スーパーカミオカンデ検出器	14
2.1.1	構造	14
2.1.2	検出原理	14
2.1.2.1	チェレンコフ光	14
2.1.2.2	光電子増倍管	15
2.2	超新星ニュートリノバースト観測	17
2.2.1	SK-Gd	17
2.2.2	SNWATCH	17
2.2.3	WIT	18
2.2.4	SK における超新星ニュートリノ	20
2.3	本研究の動機と方針	22
第 3 章	各エネルギー領域に対する事象選別	24
3.1	低エネルギー事象に対する選別	24
3.1.1	再構成の確からしさ Goodness	24
3.1.2	内部検出器壁面からの距離 $D_{\text{wall}}$	27
3.1.3	再構成エネルギー $E$	28
3.1.4	再構成事象時刻 $t$	31
3.2	高エネルギー事象に対する選別	32
第 4 章	DBSCAN fitter	36
4.1	DBSCAN	36

4.2	解析フロー	37
4.2.1	設定した閾値による事象の選別	38
4.2.2	パラメータ推定	38
4.2.3	DBSCAN によるクラスタ選別	39
4.2.4	SN 方向再構成	39
4.2.5	Estimator による結果選別	42
第 5 章	評価指標と DBSCAN-fitter におけるパラメータの最適化	44
5.1	モンテカルロシミュレーション (MC)	44
5.2	評価指標	45
5.2.1	角度分解能	45
5.2.2	Failure rate	46
5.2.3	Reconstruction time	48
5.3	DBSCAN パラメータ推定	48
5.3.1	eps の最適化 (Up_eps、Middle_eps、Bottom_eps)	48
5.3.2	各 eps に対する min_samples の推定式構築	53
5.3.3	DBSCAN パラメータ表	56
5.4	k-Nearest Neighbors 再構成法のパラメータ推定	57
5.5	Minimum Distance Estimator のパラメータ推定	62
第 6 章	結果	65
6.1	DBSCAN の性能評価	65
6.1.1	複数パラメータセット導入による安定性の向上	65
6.1.2	最終 Not-finding-rate	68
6.2	DBSCAN-fitter 後半処理の評価	68
6.2.1	kNN 再構成法における cos 閾値による角度分解能の変化	69
6.2.2	Estimator の性能評価	71
6.3	角度分解能の比較	73
6.4	Failure rate の比較	75
6.5	Reconstruction time	75
第 7 章	まとめ・今後の展望	78

## 目次

1.1	恒星の質量による最期の違い [1] .....	10
1.2	超新星爆発に至るまでの過程 [3] .....	13
2.1	スーパーカミオカンデ検出器 [5] .....	15
2.2	チェレンコフ光の伝播 .....	16
2.3	ID に設置された直径約 50cm の光電子増倍管 [6] .....	16
2.4	光電子増倍管の内部構造 (ベネシアン・ブラインド型)[7] .....	16
2.5	スーパーカミオカンデにおける WIT を用いたデータ収集システム [12] .....	19
2.6	水中で起きる超新星ニュートリノ相互作用の反応断面積 [14] .....	21
2.7	バックグラウンド事象を含めた、10kpc における中里モデルに基づく超新星ニュートリノ事象のエネルギー分布。ニュートリノ振動は考慮している。バックグラウンドによって SN ニュートリノ事象が埋もれている。 .....	22
2.8	バックグラウンド事象を含めない場合の、10kpc における中里モデルに基づく超新星ニュートリノ事象のエネルギー分布。ニュートリノ振動は考慮している。ES 事象は 10MeV 以下の低エネルギー領域に集中する。 .....	22
3.1	SN 距離 10kpc、300 個の MC データによる SN ニュートリノ事象毎に見た低エネルギー事象の Goodness 分布 (左上: バックグラウンド事象、右上:IBD 反応由来の中性子、左下:IBD 反応由来の陽電子、右下:ES 反応由来の電子)。バックグラウンド事象は 0.45 付近に、ES 事象は 0.5-0.6 にピークを示す。両者の分布は一部重なるため、閾値設定が結果に大きく影響する。 .....	25
3.2	低エネルギー事象に対して Goodness 閾値を変化させた際の Significance。いずれの SN 距離においても Goodness 閾値が 0.55 付近で最大となる。 .....	26
3.3	低エネルギー事象における ES 事象の運動方向 $\hat{u}_i$ と SN 方向 $\hat{d}_{SN,true}$ の $\cos \theta_i$ と再構成の確からしさ Goodness の 2 次元分布 (左:4kpc、中央:10kpc、右:20kpc)。いずれの SN 距離でも Goodness が 0.5-0.6 の領域に SN 方向と強い相関を持つ ES 事象が多く含まれる。 .....	27
3.4	内部検出器における座標の定義 .....	28
3.5	SN 距離 10kpc、300 個の MC データによる各 Goodness 閾値条件下での低エネルギー事象の Dwall 分布。Goodness 閾値条件下では、SN ニュートリノ事象と比べて、多数のバックグラウンド事象が ID 壁面近傍に分布している。 .....	29

3.6	低エネルギー事象における ES 事象の運動方向 $\hat{u}_i$ と SN 方向 $\hat{d}_{SN,true}$ の $\cos \theta_i$ と ID 壁面からの距離 $D_{wall}$ の 2 次元分布 (左:4kpc、中央:10kpc、右:20kpc)。いずれの SN 距離においても $D_{wall}$ が 200cm から 300cm の範囲に多くの SN 方向と強い相関のある ES 事象が分布している。 . . . . .	29
3.7	SN 距離 10kpc、300 個の MC データにおける低エネルギー事象のエネルギー分布。ほぼ全てのバックグラウンドが 4MeV 以下の領域に分布している。 . . . . .	30
3.8	低エネルギー事象における ES 事象の運動方向 $\hat{u}_i$ と SN 方向 $\hat{d}_{SN,true}$ の $\cos \theta_i$ と再構成エネルギー $E$ の 2 次元分布 (左:4kpc、中央:10kpc、右:20kpc)。いずれの SN 距離においても 3MeV 以上の領域に SN 方向と強い相関を持つ ES 事象が多く分布する。 . . . . .	30
3.9	SN 距離 10kpc、MC データ 300 個による Goodness 閾値下での低エネルギー事象の時間分布。SN ニュートリノ事象は 1 秒以内にピークを示し、バックグラウンドは一律に分布する . . . . .	31
3.10	低エネルギー事象に対して各パラメータの候補閾値を組み合わせて求めた Significance。いずれの SN 距離においても Goodness 閾値が 0.5、 $D_{wall}$ 閾値が 2m の時に Significance が高くなる。Significance が最大となる時間閾値は SN 距離によって異なる。 . . . . .	32
3.11	SN 距離 10kpc、300 個の MC データによる高エネルギー事象の Goodness 分布。ES 事象は Goodness>0.5 に分布する一方、バックグラウンド事象と IBD 由来中性子事象は Goodness<0.5 の領域にも分布する。このため、Goodness 選別が有効と示唆される。 . . . . .	33
3.12	SN 距離 10kpc、300 個の MC データによる Goodness 閾値条件下での高エネルギー事象の $D_{wall}$ 分布。低エネルギー事象と比べてバックグラウンドが少ない。 . . . . .	34
3.13	高エネルギー事象における ES 事象の運動方向 $\hat{u}_i$ と SN 方向 $\hat{d}_{SN,true}$ の $\cos \theta_i$ と ID 壁面からの距離 $D_{wall}$ の 2 次元分布 (左:4kpc、中央:10kpc、右:20kpc)。いずれの SN 距離においても ID 壁面近傍でも SN 方向と強い相関を持つ ES 事象が多く分布する。 . . . . .	35
3.14	高エネルギー事象に対する $D_{wall}$ 候補閾値を組み合わせて求めた Significance。いずれの SN 距離においても $D_{wall}$ 閾値を設定しない方が Significance が高くなる。また、Goodness 閾値 0.5 以下で安定して高い Significance を示す。 . . . . .	35
4.1	DBSCAN のアルゴリズム (ここでは min_samples を 3 にした状況を考えている) . . . . .	37
4.2	IBD タグ付けと各種事象情報に基づく選別後の再構成運動方向分布 . . . . .	38
4.3	DBSCAN による選別後の全事象運動方向分布 (赤:DBSCAN が抽出したクラスタ領域) . . . . .	40
4.4	k-Nearest Neighbors のアルゴリズム (赤点は対象のデータポイント、緑は対象のデータポイントに最も近い 3 個の近傍点、青はデータポイント間の距離を示す) . . . . .	41

4.5	全事象の運動方向 $\hat{u}_i$ と SN 方向 $\hat{d}_{SN,true}$ の $\cos \theta_i$ と MeanKdistance による散布図。 MeanKdistance が小さい事象ほど、SN 方向と強い相関を持つことが分かる。 . . . . .	42
5.1	真の SN 方向と再構成方向間の $\Delta\theta_{SN}$ . . . . .	46
5.2	SN 距離 10kpc における 919 個の MC に対し DBSCAN-fitter を適用して得た $\Delta\theta_{SN}$ 分布 . . . . .	46
5.3	20kpc の MC データ群に対して DBSCAN-fitter を適用して得た $\cos \Delta\theta_{SN}$ 分布。 $\cos \Delta\theta_{SN}=1$ 付近にピークを持ち、それ以外の領域では一様分布が見られる。 . . . . .	47
5.4	10kpc の訓練データに対し、 $\text{eps} = 0.4$ に固定して DBSCAN を適用した際の角度差 $\Delta\theta_{SN}$ ( $\text{min\_samples}$ 依存性)。良好な角度差を示す $\text{min\_samples}$ の連続領域が見られる。 . . . . .	50
5.5	10kpc の訓練データ 300 個の用いて得られた、各 $\text{eps}$ に対する「良好な角度差領域をもつデータ」数。多くの訓練データに対して良好な角度差 $\Delta\theta_{SN}$ を得やすい $\text{eps}$ 領域がある。 . . . . .	51
5.6	高エネルギー事象に対して、5.2.1 項の手順に従って抽出される平均 $\text{eps}$ と各距離に対応する平均事象数の関係 (この分布における縦線は、「良好な角度差領域を持つデータ数」が多い 10 個の $\text{eps}$ の範囲を示す) . . . . .	52
5.7	高エネルギー事象に対して、上記の手順に従って抽出される $\text{Middle\_eps}$ より高い、または低い平均 $\text{eps}$ と各距離に対応する平均事象数の関係 . . . . .	53
5.8	全エネルギー事象に対して、上記の手順に従って抽出される $\text{Middle\_eps}$ より高い、または低い平均 $\text{eps}$ と各距離に対応する平均事象数の関係 . . . . .	53
5.9	10kpc、100 個の訓練データに対し $\text{eps}$ を 0.4 に固定した場合における、各訓練データのラベルと、 $\text{middle\_eps}$ 抽出過程 2 から得られた「良好な角度差を持つ $\text{min\_samples}$ の範囲」との関係。良好な角度差を示す $\text{min\_samples}$ の値は同一距離でおおよそ一致する。 . . . . .	54
5.10	高エネルギー事象に対して、5.2.2 項の手順に従って抽出される平均 $\text{min\_samples}$ と各距離に対応する平均事象数の関係 (左:Bottom_ $\text{eps}$ に対応、中央:Middle_ $\text{eps}$ に対応、右:Up_ $\text{eps}$ に対応)。それぞれべき乗則的な関係が見られる。 . . . . .	55
5.11	全エネルギー事象に対して、5.2.2 項の手順に従って抽出される平均 $\text{min\_samples}$ と各距離に対応する平均事象数の関係 (左:Bottom_ $\text{eps}$ に対応、中央:Middle_ $\text{eps}$ に対応、右:Up_ $\text{eps}$ に対応)。それぞれべき乗則的な関係が見られる。 . . . . .	56
5.12	10kpc、300 個の訓練データにおいて、DBSCAN 分類後のクラスタに対し、k-Nearest Neighbors を適用して得られる MeanKdistance 分布。 $\text{n\_neighbors}$ が大きくなると ES 事象と非 ES 事象の集中領域が分離する。 . . . . .	58

5.13	300 個の訓練データに対して得られた <code>n_neighbors</code> と最大 Significance の分布 (左:SN 距離 10kpc、右:SN 距離 20kpc)。 <code>n_neighbors</code> の増加に伴って、Significance は単調に増加していき、最終的に飽和する。 . . . . .	58
5.14	高エネルギー事象において飽和 Significance の 95%を示す <code>n_neighbors</code> と平均事象数の関係 . . . . .	59
5.15	全エネルギー事象において飽和 Significance の 95%を示す <code>n_neighbors</code> と平均事象数の関係 . . . . .	59
5.16	10kpc において、MeanKdistance と Cosine 値の閾値を 0.9 に設定したときの Fraction 分布。青点が各ビンの Fraction を、赤線がそれらをフィッティングして得られる直線を示す。 . . . . .	60
5.17	高エネルギー事象における、抽出した重み付け関数パラメータ $a$ と平均事象数の関係 (左:cos 閾値=0.85、中央:cos 閾値=0.90、右:cos 閾値=0.95) . . . . .	60
5.18	高エネルギー事象における、抽出した重み付け関数パラメータ $b$ と平均事象数の関係 (左:cos 閾値=0.85、中央:cos 閾値=0.90、右:cos 閾値=0.95) . . . . .	61
5.19	全エネルギー事象における、抽出した重み付け関数パラメータ $a$ と平均事象数の関係 (左:cos 閾値=0.85、中央:cos 閾値=0.90、右:cos 閾値=0.95) . . . . .	61
5.20	全エネルギー事象における、抽出した重み付け関数パラメータ $b$ と平均事象数の関係 (左:cos 閾値=0.85、中央:cos 閾値=0.90、右:cos 閾値=0.95) . . . . .	61
5.21	SN 距離 10kpc の訓練データ 300 個に対する、全エネルギー事象での <code>N_MDE</code> と角度分解能の分布 . . . . .	62
5.22	SN 距離 10kpc の訓練データ 300 個に対する、全エネルギー事象での <code>N_MDE</code> と角度分解能の分布に対し隣接する 5 点の移動平均を適用した結果 . . . . .	62
5.23	高エネルギー事象において、安定な良好分解能を示す <code>N_MDE</code> の連続領域の中央値と平均事象数の関係 (左上: 単純平均、右上:kNN (cos 閾値=0.85)、左下:kNN(cos 閾値=0.90)、右下:kNN(cos 閾値=0.95)) . . . . .	64
5.24	全エネルギー事象において、安定な良好分解能を示す <code>N_MDE</code> の連続領域の中央値と平均事象数の関係 (左上: 単純平均、右上:kNN (cos 閾値=0.85)、左下:kNN(cos 閾値=0.90)、右下:kNN(cos 閾値=0.95)) . . . . .	64
6.1	高エネルギー事象において、推定される <code>min_samples</code> を基準とし、その値を一定割合だけ増減させた二種類の <code>min_samples</code> を追加した場合の Not-finding-rate の比較。横軸は追加した <code>min_samples</code> が推定値に対して何%に相当するかを示す) . . . . .	67
6.2	全エネルギー事象において、推定される <code>min_samples</code> を基準とし、その値を一定割合だけ増減させた二種類の <code>min_samples</code> を追加した場合の Not-finding-rate の比較。横軸は追加した <code>min_samples</code> が推定値に対して何%に相当するかを示す . . . . .	67

6.3	高エネルギー事象および全エネルギー事象に対する最終的な Not-finding-rate。両者とも 20kpc におけるクラスタ抽出の失敗率が 2.5%程度となった。 . . . . .	69
6.4	高エネルギー事象に対して DBSCAN-fitter を適用した場合における各 cos 閾値の角度分解能。単純平均による再構成法 (青線) よりも、cos 閾値=0.95 での kNN 再構成法 (赤線) の方がより良い分解能を示す。 . . . . .	71
6.5	全エネルギー事象に対して DBSCAN-fitter を適用した場合における各 cos 閾値の角度分解能。単純平均による再構成法 (青線) よりも、cos 閾値=0.95 での kNN 再構成法 (赤線) の方がより良い分解能を示す。 . . . . .	71
6.6	高エネルギー事象と全エネルギー事象における Estimator Failure-rate。青線が高エネルギー事象による結果、オレンジ線が全エネルギー事象による結果を示す . . . . .	72
6.7	最良の角度差 $\Delta\theta_{SN,best}$ と MDE によって選択された結果の角度差 $\Delta\theta_{SN,MDE}$ との関係を示す散布図 (左: 高エネルギー事象、右: 全エネルギー事象)。MDE は多くの SN 事象に対して最良の角度差を持った結果を抽出できていないことを示している。 . . . . .	73
6.8	DBSCAN-fitter と現行の SNWATCH 方向再構成法による最終的な角度分解能の比較	75
6.9	DBSCAN-fitter と現行の SNWATCH 方向再構成法による最終的な Failure-rate の比較	76

# 第 1 章 重力崩壊型超新星爆発

超新星爆発 (Supernova explosion: SNe) は、恒星がその進化の最終段階において引き起こす爆発現象である。特に、大質量星のコアが重力不安定に陥ることで生じる爆発は、重力崩壊型超新星爆発 (Core-collapse Supernova Explosion: CCSNe) と呼ばれる。CCSNe は銀河系内でおよそ数十年から 100 年に一度の頻度で発生すると考えられており、その発生は天文学的に極めて稀である一方で爆発時には膨大なエネルギーが短時間に放出される。これまでに人類が観測した超新星ニュートリノバーストは、1987 年に大マゼラン雲で発生した SN1987A に由来するもののみであり、当時のカミオカンデ検出器で約 12 秒間に 11 事象が観測された。この観測により、超新星爆発機構およびニュートリノの性質に関する重要な知見が得られた [1, 2, 3]。このことは、超新星爆発においてニュートリノが爆発初期の情報を直接担う粒子であることを示しており、ニュートリノ観測は爆発機構のみならず、超新星爆発の早期検知や方向決定においても重要な役割を果たすことを示唆している。

## 1.1 爆発過程

本節では、重力崩壊型超新星爆発 (CCSNe) の発生過程について概説する。CCSNe は大質量星の進化の終末に生じる重力崩壊現象であり、中心部の崩壊、衝撃波の形成・停滞・復活といった複数の過程を経て爆発に至る。以下ではその物理過程の概要を述べる。

### 1.1.1 星の進化

恒星は自己重力と、内部での核融合反応によって生じる圧力との釣り合いによって力学的平衡を保っている。大質量星  $\geq 8M_{\odot}$  ( $M_{\odot}$ : 太陽質量) では、核融合反応の進行に伴い中心部に重元素が蓄積され、最終的に鉄を主成分とするコアが形成される。鉄はそれ以上の核融合によってエネルギーを生み出すことはできないため、コア質量は増大し続ける。この段階において、鉄コアは主として電子の縮退圧によって支持されている。鉄コアの質量が Chandrasekhar 質量 ( $\sim 1.4M_{\odot}$ ) に達すると、電子の縮退圧では自己重力を支えきれなくなり、急激な重力崩壊が開始される。この重力崩壊を基点として生じる超新星爆発が、重力崩壊型超新星爆発 (CCSNe) であり、その結果として、中心天体には中性子星、あるいはブラックホールが形成される。図 1.1 に恒星の質量による最期の違いを示す。

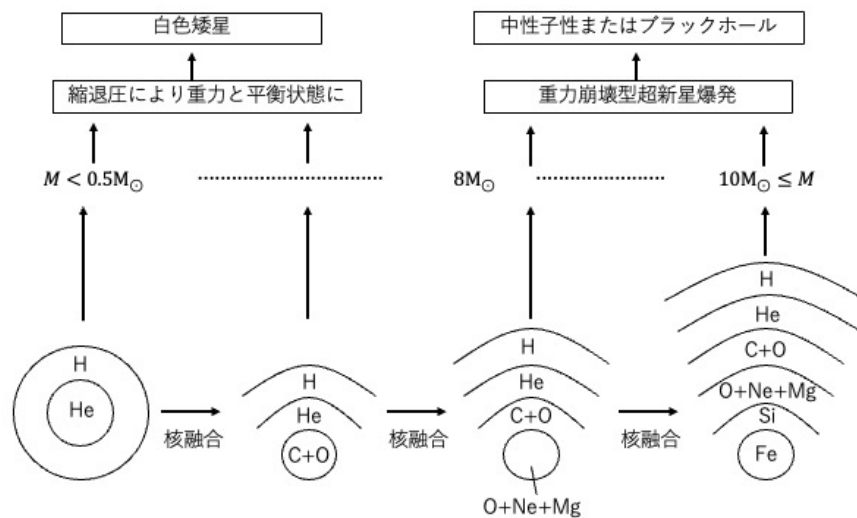


図 1.1: 恒星の質量による最期の違い [1]

### 1.1.2 超新星の分類

超新星は、観測されるスペクトルおよび明るさの時間変化を表す光度曲線に基づいて分類される。スペクトル中に水素線を示さない方を I 型、水素線を示す方を II 型と呼ぶ。I 型のうち、ケイ素線を示すものは Ia 型に分類される。

現在では、Ia 型超新星は主に熱核反応によって引き起こされる一方、それ以外の多くの型の超新星は大質量星の重力崩壊によって生じると理解されている。本研究では、重力崩壊型超新星爆発に伴うニュートリノ放出およびその観測情報を用いた解析を対象とするため、以降は CCSNe のみを扱う。

### 1.1.3 重力崩壊

質量が  $8M_{\odot} \leq M \leq 10M_{\odot}$  の恒星では、核融合反応の進行により酸素・ネオン・マグネシウムを主成分とする高密度なコアが形成される。このような環境下では、以下に示す電子捕獲反応が進行する。



電子捕獲反応によりコア内部の電子数が減少すると、電子縮退圧が低下し、自己重力に支えきれなくなることで重力圧縮が進行する。

一方、初期質量が  $\geq 10M_{\odot}$  の恒星では、核融合反応は最終的に鉄の合成で停止し、鉄を主成分とするコアが形成される。鉄はそれ以上の核融合によってエネルギーを生成できないため、コア質量は増加し続ける。中心温度が約  $5 \times 10^9$  K に達すると、高エネルギー光子による以下

の光分解反応が起こる。



これらはいずれも吸熱反応であり、コア内の熱エネルギーを奪うことで圧力も低下させ、重力崩壊をさらに加速させる。このように高エネルギー光子によって原子核が分解される過程を光分解という。

### 1.1.4 ニュートリノトラッピング

重力崩壊の進行に伴い、電子捕獲反応によってコア内部の中性子数は増加し、密度は急激に上昇する。密度が  $\rho \sim 10^{12} \text{ g/cm}^3$  に達すると、生成された電子ニュートリノは以下に示す原子核との中性散乱反応を頻繁に起こすようになる。



この結果、ニュートリノの平均自由行程は著しく短くなり、コア内部から脱出できなくなる。この現象をニュートリノトラッピングと呼ぶ。ニュートリノが実質的に閉じ込められる領域はニュートリノ球と呼ばれ、その外側からのみニュートリノは自由に放出される。このトラッピングの存在により、重力崩壊の後半段階において大量のニュートリノが短時間に放出される条件が整う。

### 1.1.5 衝撃波の発生

鉄の生成が進行すると、星の中心部の密度は急激に増加し、収縮速度の違いによりコアは内部コアと外部コアに分離する。内部コアは亜音速で収縮する一方、外部コアは超音速で収縮する。

中心密度が原子核密度 ( $\rho \sim 10^{14} \text{ g/cm}^3$ ) に達すると、電子捕獲反応によって生成された中性子による縮退圧が支配的となり、内部コアの収縮は停止する。さらに、この高密度状態における圧力反発により、内部コアは急激に膨張へと転じる。この現象は反跳 (バウンス) と呼ばれる。一方、外部コアは依然として超音速で収縮を続けているため、両者の境界において衝撃波が形成される。形成された衝撃波は当初外向きに伝播するが、主に以下の二つの要因により、次第にその運動エネルギーを失い停滞する。

- 衝撃波によって加熱された物質中で光分解反応が促進され、吸熱反応によって衝撃波のエネルギーが消費されること。
- 衝撃波通過後の高密度領域において電子捕獲反応が進行し、多数のニュートリノが生成・放出され、エネルギーが持ち去られること。

これらの効果により衝撃波は一時的に停滞する。その後、ニュートリノ加熱や磁場の増幅などの効果によって、衝撃波は数 100ms 後に再び加速されると考えられている。衝撃波が恒星表面に到達すると、星全体が膨張し、可視光で明るく観測される超新星爆発が生じる。

### 1.1.6 中性子化バースト

衝撃波が発生すると、式 (1.2) に示した光分解反応が活発になり、ニュートリノを閉じ込めていた原子核は核子へと分解される。核子上での電子捕獲反応率は原子核に比べてはるかに高いため、電子捕獲反応が急速に進行し、大量の電子ニュートリノが生成される。衝撃波がニュートリノ球を超えると、生成された電子ニュートリノは外部へ一斉に放出される。この短時間に生じる電子ニュートリノの放出を中性子化バーストと呼ぶ。中性子化バーストは数 ms の短時間に渡って発生し、その間に放出されるエネルギーは  $\leq 10^{51}$  erg に達する。これは、超新星爆発全体に伴うニュートリノ放出エネルギー ( $\sim 10^{53}$  erg) の約 1% に相当する。中性子化バースト終了後も、形成された原始中性子星の冷却に伴い、数秒間にわたって準定常的なニュートリノ放出が継続する。中性子化バーストでは、主に電子捕獲反応に由来する電子ニュートリノが支配的であるが、その後の冷却段階では、電子型ニュートリノおよび反ニュートリノに加え、ミュー型・タウ型ニュートリノを含む全フレーバーのニュートリノがほぼ同程度放出される。これらのニュートリノは以下の反応によっても生成される。以下に示す反応式中の  $\nu$  は、電子型、ミュー型、タウ型のいずれかのニュートリノを表す。

$$e^- + e^+ \rightarrow \nu + \bar{\nu} \quad (1.5)$$

$$N + N \rightarrow N + N + \nu + \bar{\nu} \quad (1.6)$$

超新星爆発に至るまでの一連の過程を 1.2 に示す。図 1.2 において、左上は重力崩壊の開始、右上は内部コアと外部コアの境界での衝撃波形成、中央左は衝撃波の停滞、中央右は衝撃波の復活、下段は衝撃波の伝播および中性子化バーストの発生過程を表している。

## 1.2 マルチメッセンジャー天文学とスーパーカミオカンデ

超新星爆発に伴って放出されるニュートリノは、爆発初期の情報を担う粒子として、マルチメッセンジャー天文学において極めて重要な役割を果たす。超新星爆発では、重力崩壊により解放される重力エネルギーの大部分 (99%) がニュートリノとして放出されることが知られている。このため、超新星ニュートリノの観測は、超新星の進化過程や爆発機構の理解に直接的な手がかりを与える。また、ニュートリノバースト観測は、超新星爆発に伴う可視光放射に先立つ早期検知情報を提供する。恒星内部で生成されたニュートリノは、重力崩壊直後に形成されるニュートリノ球の外側から放出され、密度の低い外層領域を光速に近い速度でほぼ直進して伝播する。一方、爆発光の起源となる衝撃波はニュートリノに比べて伝播速度が著しく遅く

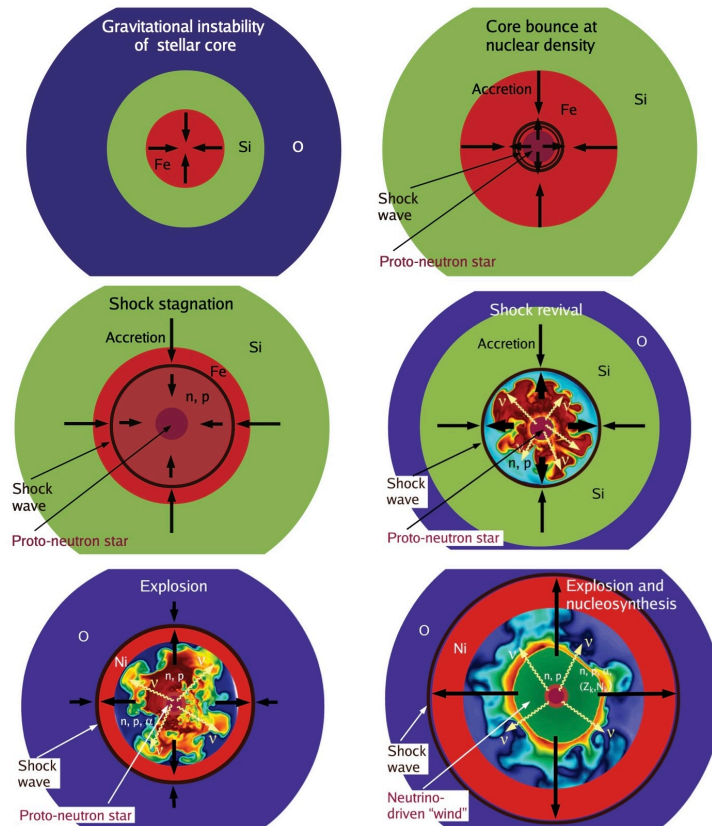


図 1.2: 超新星爆発に至るまでの過程 [3]

( $\sim 10^4$ km/s)、恒星表面に到達するまでに長い時間を要する。この時間差は恒星の半径や外層の密度構造に依存し、数十分から数日に及ぶ。したがって、超新星ニュートリノバーストの観測は、可視光が地球に到来するのに先立って超新星爆発を検知できるため、早期アラートの発信を通じて爆発光の光学観測や重力波観測の初動対応を可能にする。特に到来方向の推定が可能であれば望遠鏡による追観測領域を大幅に絞り込むことができるため、マルチメッセンジャー観測の成功率の向上に寄与する。このように超新星ニュートリノ観測は爆発機構の解明において重要な位置付けを持つ。さらに、観測機会が極めて限られることから、世界各地の実験において観測体制が整備されている。

現在、SN1987A 由来の超新星ニュートリノバーストを観測したカミオカンデを発展させたスーパーカミオカンデでは、銀河系内で超新星爆発が起きた場合、数千規模の超新星ニュートリノ事象が観測されると予想されている。観測される事象のうち、特に電子散乱反応に由来する事象が超新星ニュートリノの到来方向と強い相関を持つことから、これを手がかりとした方向再構成手法が導入されている。さらに得られた方向情報は望遠鏡による追観測にも有用であり、スーパーカミオカンデでは早期にアラートを発信する体制が整えられている。第 2 章では、スーパーカミオカンデにおける超新星ニュートリノバースト観測の詳細について述べる。

# 第2章 スーパーカミオカンデにおける超新星ニュートリノバースト観測

## 2.1 スーパーカミオカンデ検出器

### 2.1.1 構造

スーパーカミオカンデ (Super-Kamiokande: SK) は、岐阜県飛騨市神岡町にある池ノ山の地下約 1000m に建造された、水チェレンコフ型の宇宙素粒子観測装置である [4]。当初の目的である陽子崩壊の探索に加え、現在ではニュートリノの性質解明や、ニュートリノを利用した宇宙観測を主な目的としている。

SK は、約 5 万トンの高純度純水を貯蔵する高さ 41.1m、直径 39.3m の円筒形水タンクから構成され、内部検出器 (Inner Detector:ID) と外部検出器 (Outer Detector:OD) に分けられている。ID は直径 33.8m、高さ 36.2m の円筒構造を持ち、約 32500 トンの超純水を貯蔵している。その内壁には直径約 50cm の光電子増倍管 (photomultiplier tube:PMT) が 11129 本設置されている。一方、OD は ID を取り囲む構造をしており、約 17500 トンの超純水を貯蔵している。OD 内部には、直径約 20cm の PMT が 1885 本配置されている。ID は水中でのニュートリノ相互作用により生成された荷電粒子が放射するチェレンコフ光を検出する役割を担い、OD は主に宇宙線ミューオンや SK 周囲の岩盤からの放射性バックグラウンド事象を識別・除去するために用いられる。図 2.1 に SK の構造を示す。

### 2.1.2 検出原理

#### 2.1.2.1 チェレンコフ光

ニュートリノは電氣的に中性で物質との相互作用が極めて弱いため、直接観測することはできない。そのため SK では、ニュートリノが水中の電子や原子核と弱い相互作用を起こして生成する荷電粒子を検出し、ニュートリノ事象を間接的に観測する。超新星ニュートリノと水標的との代表的な反応として電子散乱反応 (式 2.1) および逆ベータ崩壊反応 (2.2) が挙げられる。

$$\nu + e^- \rightarrow \nu + e^-, \quad (2.1)$$

$$\bar{\nu}_e + p \rightarrow e^+ + n. \quad (2.2)$$

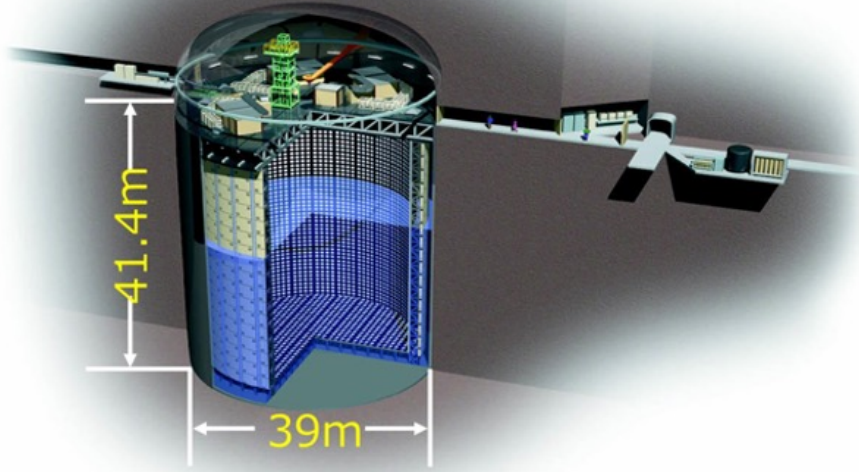


図 2.1: スーパーカミオカンデ検出器 [5]

ここで  $\nu$  は電子ニュートリノ, ミューニュートリノ, またはタウニュートリノのいずれかを,  $\bar{\nu}_e$  は反電子ニュートリノを表す。

生成された荷電粒子が媒質中 (SK では水) を媒質中の光速度より速く運動する場合、媒質の分極が緩和される過程でチェレンコフ光が放射される。図 2.2 にその概念図を示す。チェレンコフ光は荷電粒子の進行方向に対して円錐状に放射され、その放射角  $\theta_c$  は媒質の屈折率  $n$  と粒子の速度  $v = \beta c$  を用いて

$$\cos \theta_c = 1/n\beta \quad (2.3)$$

と表される。水の屈折率は  $n \simeq 1.33$  であり、 $\beta = 1$  の場合  $\theta_c \simeq 42^\circ$  となる。

### 2.1.2.2 光電子増倍管

光電子増倍管 (PhotoMultiplier Tube: PMT) は、入射した光を電気信号に変換する光検出器である。SK の内部検出器内壁には直径約 50cm の PMT が設置されており、その外観を図 2.3 に、内部構造を図 2.4 に示す。PMT による光検出は以下の過程に従って行われる。

1. チェレンコフ光が PMT 前面の光電面 (バイアルカリ:Sb-K-Cs) に入射すると、光電効果により光電子が生成される。
2. 生成された光電子が集束電極によって電子増倍部へと導かれる。
3. 光電子は高電圧によって加速され、複数段のダイノードで二次電子放出を繰り返すことで増幅される。
4. 増幅された電子はアノードに集められ、初期の光電子数に対して約  $10^6$  倍の電流信号とし

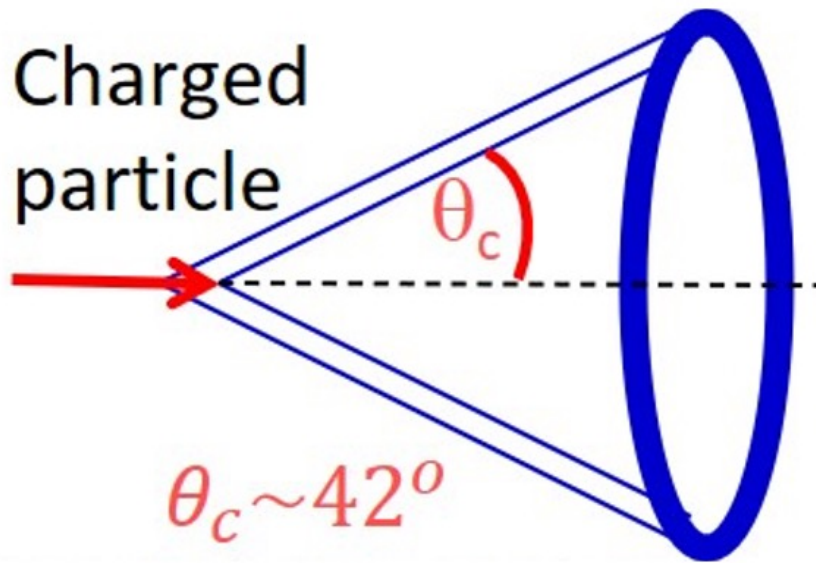


図 2.2: チェレンコフ光の伝播

て出力される。



図 2.3: ID に設置された直径約 50cm の光電子増倍管 [6]

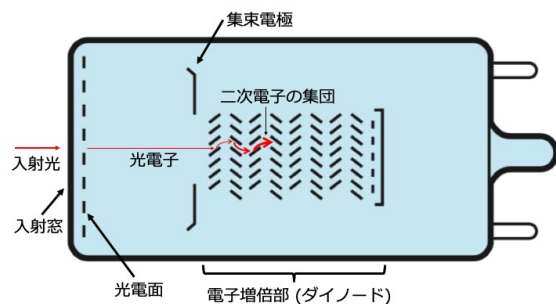


図 2.4: 光電子増倍管の内部構造 (ベネシアン・ブラインド型)[7]

ニュートリノ相互作用によって生成される荷電粒子が放出するチェレンコフ光は、ID 内に設置された PMT によって検出される。検出された PMT 信号の到達時刻、光量、および PMT 位置情報を用いることで、事象の発生位置 (vertex)、エネルギー、および荷電粒子の運動方向を再構成することができる。本研究では Wide-band Intelligent Trigger(WIT[8]) によって取得・再構成された事象情報を解析対象とする。WIT による再構成手法および本研究で用いる事象情報については、後の節で述べる。

## 2.2 超新星ニュートリノバースト観測

本節では、スーパーカミオカンデにおける超新星ニュートリノバースト観測に向けた取り組みや観測時の対応について述べる。

### 2.2.1 SK-Gd

スーパーカミオカンデにおいて超新星ニュートリノバースト観測から SN 方向を推定する際、検出される事象の大部分は逆ベータ崩壊 (Inverse Beta Decay:IBD, 式 2.2) 由来である。IBD によって生成される陽電子と中性子の運動方向は入射ニュートリノ方向とほとんど相関を持たない。そのため、SN 方向推定において IBD 事象は、方向感度を持つ電子散乱 (Elastic Scattering:ES, 式 2.1) 事象に対する主要なバックグラウンドとなる。このバックグラウンドを低減するために、SK では IBD に伴って生成される中性子を識別・タグ付けすることを目的として、超純水にガドリニウムを溶解する SK-Gd 実験 [9] が 2020 年夏より開始された。

純水中では、中性子は平均捕獲時間  $200\mu\text{s}$  程度で水素原子核に捕獲され、 $2.2\text{MeV}$  のガンマ線を放出する。しかしこのエネルギー領域では低エネルギーバックグラウンド事象との識別が難しく、中性子検出効率も約 25%と低い。

一方、ガドリニウムは中性子反応断面積が大きく、中性子捕獲時に合計約  $8\text{MeV}$  のガンマ線を放出する。そのため、純水の場合に比べて遅延信号 (delayed) の検出が容易になる。IBD 事象の識別は、陽電子による即時信号 (prompt) と、中性子が Gd に捕獲されて生成されるガンマ線に由来する遅延信号 (delayed) との時間相関を利用した同時遅延計測により行われる。

SK-Gd 実験では、ガドリニウム硫酸塩八水和物  $\text{Gd}_2(\text{SO}_4)_3 \cdot 8\text{H}_2\text{O}$  を溶解しており、現在の運用ではガドリニウム濃度は質量比で 0.033%である。この条件下での中性子捕獲効率は約 75%[10]である。

### 2.2.2 SNWATCH

スーパーカミオカンデには超新星爆発由来のニュートリノバーストをリアルタイムで監視しつつ、検出した際にはアラートを発信するシステムがあり、これを Real-time SN neutrino burst monitor(SNWATCH[11]) と呼ぶ。

SK 内で超新星ニュートリノ相互作用が起これると、生成される荷電粒子が水中でチェレンコフ光を放射し、その光は PMT によって電気信号として検出される。検出された事象はオンライン再構成プログラムを用いて、事象発生位置 (vertex)、運動方向およびエネルギーへと再構成される。SK の標準再構成プログラムではこの処理に約 1 分を要するのに対し、WIT では約 25 秒で完了する (WIT の詳細は後述する)。再構成後、ガドリニウムを用いた IBD 事象タグ付けにより IBD 事象を識別し、方向推定におけるバックグラウンドとして除去する。さらに、

検出器由来バックグラウンドの多い 6MeV 以下の事象や宇宙線ミュオンを除去する。

SNWATCH におけるアラート発信基準として、20 秒の時間窓内に存在する選別後イベント数 ( $N_{cluster}$ ) と、再構成された vertex 分布の形状を示す指標  $D$  を用いる。 $D$  は 0 から 3 の整数値をとり、 $D = 0$  は点状分布、 $D = 1$  は線状分布、 $D = 2$  は平面状分布、 $D = 3$  は体積状分布を示す。超新星ニュートリノは検出器タンク全体でほぼ一様に相互作用するため、バースト観測時であれば  $D = 3$  が期待される。

$D = 3$  かつ  $N_{cluster} \geq 60$  を満たした場合、超新星爆発発生した可能性が極めて高いと判断され、NASA の General Coordinates Network(GCN) へ自動通知される (Golden Alarm)。一方、 $N_{cluster}$  が 60 に達しない場合でも、 $N_{cluster} \geq 25$ , かつ  $D = 3$  を満たす場合には、expert 向けに通知が送られるとともに、Supernova Early Warning System(SNEWS) にも通知される (Normal Warning)。

現在の運用では、超新星バースト由来の事象を観測してから約 90 秒以内にアラートを発信することが可能であり、その際には観測された事象数、検出器の稼働状況、並びに再構成された超新星爆発の方向情報などが報告される。このように SNWATCH はアラート発信のみならず、超新星爆発の方向推定によって追観測領域を絞り込み、マルチメッセンジャー観測の初動対応を最適化する上で重要な役割を担う。

一方で、現行の SNWATCH では方向推定に有効な電子散乱事象が多く含まれる低エネルギー領域の事象が十分に活用されていない。そのため、低エネルギー領域を活用した方向推定手法の開発には改善の余地がある。

### 2.2.3 WIT

超新星ニュートリノバーストで観測されるニュートリノのエネルギーは概ね 100MeV 以下であり、このエネルギー領域では各 PMT が検出するチェレンコフ光は単一光子に対応することが多い。そのため、電荷情報は再構成において有効に機能しにくく、PMT ヒットの到達時刻情報が事象再構成において重要な役割を果たす。SK における PMT の時間分解能は数 ns 程度であり、検出器が大規模であることから、チェレンコフ光が検出器を横断するまで最大約 230ns を要する。

10MeV 以上のエネルギーでは十分なチェレンコフ光量が得られるため、比較的安定した事象再構成が可能である。一方、6MeV 以下では、チェレンコフヒット数が 25-30 個程度まで減少し、さらにエネルギーが低下するにつれて放射性バックグラウンドが急増する。このため、低エネルギー事象を信頼性高く再構成するためには、高度なオンライン再構成アルゴリズムが必要となる。

この低エネルギー領域の事象を対象として導入されたのが、Wide-band Intelligent Trigger(WIT) システムである [8, 12]。WIT は、生データをリアルタイムで処理し、反跳電子の運動エネルギー約 2.5MeV までの事象をほぼ 100%の効率で抽出可能なオンライン・トリ

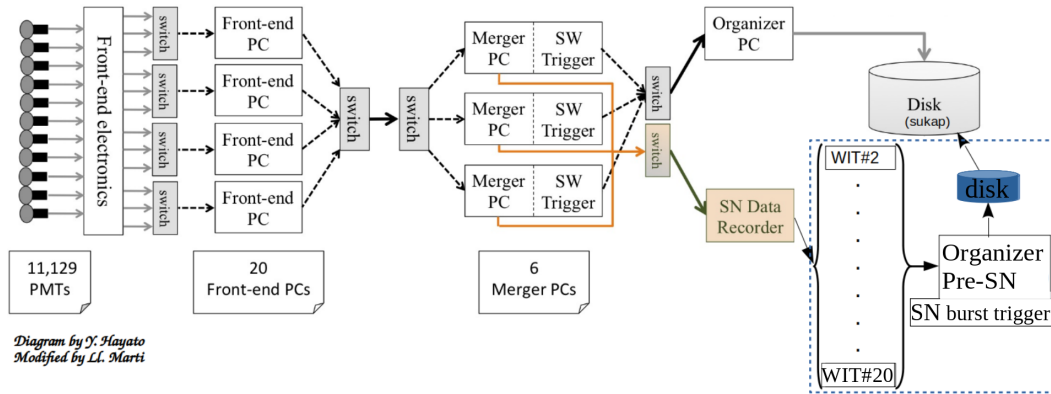


図 2.5: スーパーカミオカンデにおける WIT を用いたデータ収集システム [12]

ガー/再構成システムである。WIT は 18 台のオンライン計算機を用いた並列計算により、大量の PMT 信号に対する高速な事象選別と再構成を実現している。図 2.5 に SK における WIT を用いたデータ収集システムフローを示す。PMT で検出された信号は、データ収集コンピュータ (Merger PCs) により約 23ms 分のデータを 1 ブロックとして統合される。その後、これらのデータは SN ニュートリノバーストを監視する SN Data Recorder を経由して WIT システムに送信され、トリガーおよび事象再構成がオンラインで実行される。

WIT における事象選別および再構成は以下の手順で行われる。まず、230ns の時間窓内に 11 ヒット以上という事前トリガー条件を課し、孤立した暗電流ノイズなどの非物理的なヒットを除去する。この処理により、バックグラウンド事象の約 70%が除去される。次に、STORE(Software Triggered Online Reconstruction of Events) フィルタを適用する。通常チェレンコフ光は電子の移動に伴って連続的に放出されるが、20MeV 以下の電子は水中での飛跡長が数 cm 程度と短く、SK における vertex 分解能と比べて十分に小さい。そのため、チェレンコフ光は点光源からの放射と近似できる。STORE ではこの仮定に基づき、PMT ヒットの到達時刻を用いて、「任意の 2 つの PMT ヒットの時間差が水中における光速での移動時間以下である」という条件を満たすヒット集合を点光源由来とみなして選別する。さらに、選別されたヒットから 4 ヒットの組み合わせを構成し、各ヒットの時間残差が全てゼロとなる条件を用いて vertex 候補を求める。ここで時間残差は

$$t_{res} = t_i - \frac{|\vec{x}_i - \vec{v}|}{c} - t_0 \quad (2.4)$$

で定義される。 $t_i$  は PMT $i$  のヒット時刻、 $\vec{x}_i$  は PMT 位置、 $\vec{v}$  および  $t_0$  はそれぞれ vertex 仮説と発生時刻仮説である。STORE 後に得られるヒット集合には反射光・散乱光・暗電流などが依然として含まれるため、高速発生点再構成アルゴリズム Clusfit を適用して孤立ヒットを除去する。最後に、高精度発生点再構成アルゴリズム BONSAI(Branch Optimization Navigating Successive Annealing Iterations) を適用する [13]。BONSAI はトリガー時刻の前後に定義された時間窓内における PMT ヒットの時間残差に対して尤度を計算し、尤度が最大と

なる仮説を再構成 vertex として採用する。

さらに、vertex 再構成の確からしさを評価する指標として Goodness を算出する。例えば、vertex 仮説  $\vec{v}$  に対する Goodness は、

$$g(\vec{v}) = \sum_i w_i \exp \left[ -\frac{1}{2} \left( \frac{t_i - |\vec{x}_i - \vec{v}|/c}{\sigma} \right)^2 \right] \quad (2.5)$$

で表される。ここで、 $w_i$  は時間残差に基づくヒット重み、 $\sigma$  はチェレンコフ事象に期待される有効時間分解能を表す。Goodness は 0 から 1 の値を取り、1 に近いほど再構成の質が高いとみなす。Goodness が低い事象や内部検出器壁面近傍の事象はオンライン段階で除去される。以上のような多段階再構成システムにより、低エネルギー領域におけるトリガーと再構成を効率的に実現するとともに、処理時間およびデータサイズの大幅な削減が可能となっている。

また WIT の導入により、超新星ニュートリノバースト観測において、従来の SNWATCH では十分に活用されていなかった低エネルギー領域の事象を解析対象に含めることが可能となる。そのため、低エネルギー領域に存在する電子散乱事象の情報を有効に活用することで、方向再構成性能の向上につながる可能性がある。

## 2.2.4 SK における超新星ニュートリノ

SK において、超新星ニュートリノと水分子が引き起こす主な相互作用として、逆ベータ崩壊 (Inverse Beta Decay:IBD)、電子散乱反応 (Elastic Scattering:ES)、酸素との荷電カレント反応 (Charged-current:O16CC)、酸素との中性カレント反応 (Neutral-current:O16NC) が挙げられる。これらの相互作用のうち、超新星爆発の方向再構成において最も重要なのは ES 反応である。ES 反応では、生成される電子が入射ニュートリノの運動方向に強く前方散乱されるため、ES 事象は SN ニュートリノの到来方向と強い相関を持つ。一方、IBD 反応や酸素原子核との相互作用では、反応によって生成される荷電粒子の運動方向は概ね等方的であり、方向再構成においてはバックグラウンドとして振る舞う。そのため、ES 事象を効率よく抽出することが、SN の方向を高精度に推定する上で重要な課題となる。

図 2.6 にこれらの反応断面積を示す。この分布から各反応の寄与がエネルギー領域によって変化することが分かる。SN ニュートリノのエネルギーは数 MeV-数十 MeV に分布し、特に 70MeV 以下の広いエネルギー領域において IBD 反応が最も大きな断面積を持つ。中でも、30MeV 以下では IBD 反応が支配的であり、観測される SN ニュートリノ事象の大部分を占める。一方、ニュートリノエネルギーが 20MeV を超える領域では、O16CC および O16NC 反応の断面積が ES 反応を上回り、エネルギーの増加とともに急速に増大する。

図 2.7 はバックグラウンド事象を含めた際の各相互作用による事象のエネルギー分布、図 2.8 はバックグラウンド事象を含めなかった際のエネルギー分布を示す。また表 2.1 に SN ニュートリノ相互作用の反応割合を示す。図 2.8 によると、SN ニュートリノ相互作用の約

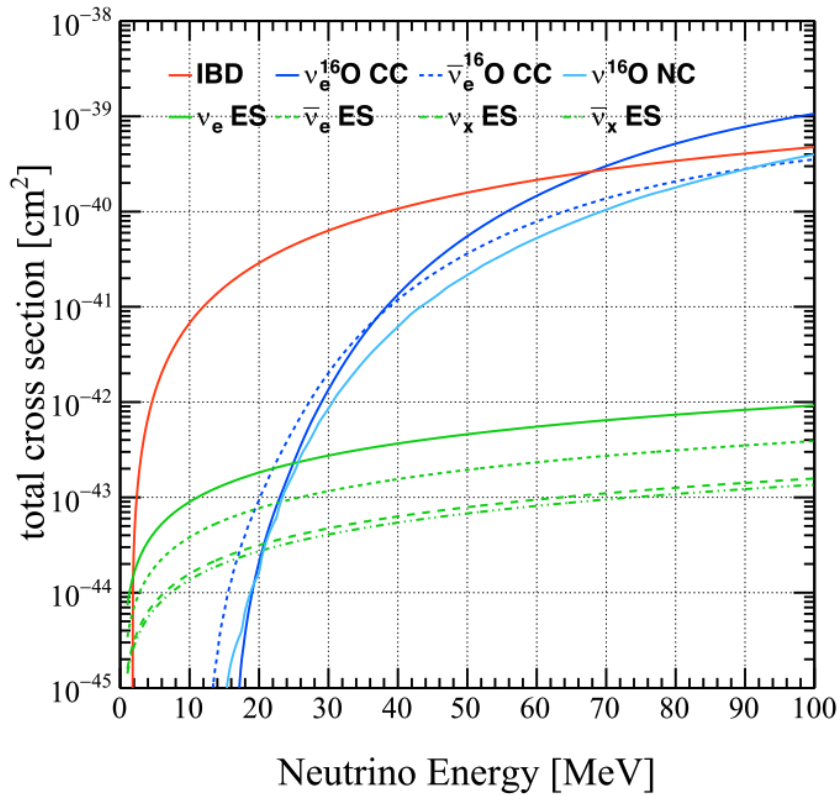


図 2.6: 水中で起きる超新星ニュートリノ相互作用の反応断面積 [14]

90%を占める IBD 事象は 10-60MeV に渡って支配的である。一方、ES 事象は全体の数%程度であり、特に 10MeV 以下の領域に多く分布し、エネルギーが増加するにつれて事象数は急速に減少する。さらに、図 2.7 によると、5MeV 以下の低エネルギー領域では検出器由来バックグラウンドが多く、信号事象の識別が困難となる。従って、ES 事象の多くが 10MeV 以下のエネルギー領域に存在することは SN 方向推定において重要である一方で、同時にバックグラウンド抑制が不可欠であることも意味する。

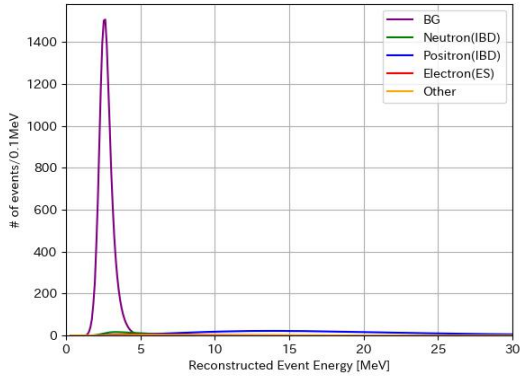


図 2.7: バックグラウンド事象を含めた、10kpc における中里モデルに基づく超新星ニュートリノ事象のエネルギー分布。ニュートリノ振動は考慮している。バックグラウンドによって SN ニュートリノ事象が埋もれている。

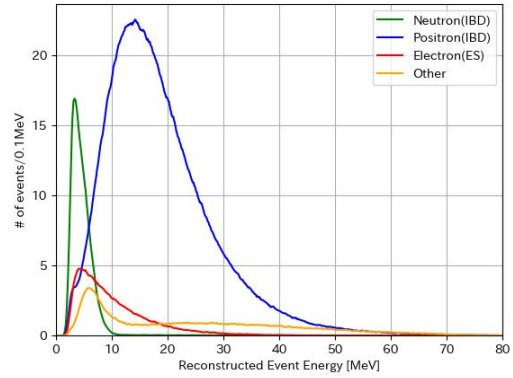


図 2.8: バックグラウンド事象を含めない場合の、10kpc における中里モデルに基づく超新星ニュートリノ事象のエネルギー分布。ニュートリノ振動は考慮している。ES 事象は 10MeV 以下の低エネルギー領域に集中する。

表 2.1: 超新星ニュートリノと SK 内の水分子との相互作用 (反応割合: 左はニュートリノ振動を考慮、右はニュートリノ振動を考慮しない場合)[15]

ニュートリノ反応式	反応割合
$\bar{\nu}_e + p \rightarrow n + e^+$	88% / 89%
$\nu_e + e^- \rightarrow \nu_e + e^-$	1.5% / 1.5%
$\bar{\nu}_e + e^- \rightarrow \bar{\nu}_e + e^-$	< 1% / < 1%
$\nu_x + e^- \rightarrow \nu_x + e^-$	1% / 1%
$\nu_e + {}^{16}\text{O} \rightarrow e^- + {}^{16}\text{F}$	2.5% / < 1%
$\bar{\nu}_e + {}^{16}\text{O} \rightarrow e^+ + {}^{16}\text{N}$	1.5% / 1%
$\nu_x + {}^{16}\text{O} \rightarrow \nu_x + O^*/N^* + \gamma$	5% / 6%

SN 方向再構成において有効な ES 事象をいかに抽出・活用するかが重要である。特に、ES 事象が多く含まれる低エネルギー領域では検出器由来バックグラウンドが増大するため、信頼性の高い事象再構成・選別手法が求められる。次節では、現行の SNWATCH の方向再構成手法の課題を整理し、低エネルギー事象の有効利用を目的とした本研究の方針について述べる。

## 2.3 本研究の動機と方針

スーパーカミオカンデでは、SNWATCH システムによって超新星ニュートリノバーストの検出と方向推定が実施されている。現行手法では、方向推定においてバックグラウンドとなる IBD 事象に対しガドリニウムを用いたタグ付けを行い、これらを除去することで ES 事象の寄与を高めている。また、検出器由来バックグラウンドに対しては 6MeV のエネルギー閾値を設定することで抑制を行っている。その後、分布全体に対する最大尤度法に基づくフィッティング手法 (Maximum-likelihood fitter: ML-fitter) を用いて SN 方向を再構成をしている。

ML-fitter は統計的に妥当な推定法である一方、尤度関数の構成には事象の角度・エネルギー分布に関するモデル仮定が含まれる。そのため、背景事象の混入や低統計条件、反応割合、エネルギー分布の変化などにより、観測される事象分布が想定モデルからずれた場合には、推定が不安定となる可能性がある。また、現行の SNWATCH では、検出器由来バックグラウンドが増加する低エネルギー領域 (6MeV 以下) の事象を解析から除去しているが、この領域には SN 方向と強い相関を持つ ES 事象も含まれるため、方向再構成に有効な情報が十分に活用されていない可能性がある。この課題に対し、WIT システムは、約 2.5MeV までのエネルギー事象を対象とした高速な事象選別と再構成を可能とし、従来よりも低エネルギー領域の情報を活用した SN ニュートリノバースト解析を実現できる可能性がある。

そこで本研究では、WIT を活用し、低エネルギー事象を含めた解析においても安定して ES 事象が形成する局所的な高密度領域を抽出することを目的として、密度に基づくクラスタリング手法 DBSCAN を導入し、この領域を直接抽出するアプローチを採用する。これは、従来のフィッティング手法とは異なり、SN 方向近傍に現れる局所的な密度構造を手がかりとして方向再構成を行う点に特徴がある。本研究では、DBSCAN を用いた方向再構成手法 (DBSCAN-fitter) を構築し、その有効性を角度分解能、Failure-rate、および再構成時間の観点から評価する。

## 第3章 各エネルギー領域に対する事象選別

2.3 節で述べたように、現行の SNWATCH では、放射性バックグラウンドや暗電流ノイズなどの検出器由来バックグラウンドの増加を避けるため、6MeV のエネルギー閾値を設定して超新星爆発の方向を再構成している。しかし、6MeV 以下の低エネルギー領域には超新星ニュートリノ到来方向と強い相関を持つ ES 事象が多く含まれており、この領域を除外することは方向再構成に有効な情報を損なう可能性がある。

一方で、2.2.4 項で述べたように、低エネルギー領域では検出器由来バックグラウンドが支配的となるため、単にエネルギー閾値を緩和して事象を追加すると、等方的に分布するバックグラウンド事象が方向推定を劣化させるだけでなく、後段の DBSCAN によるクラスタ抽出においても誤ったクラスタの形成や抽出失敗を引き起こし得る。従って、低エネルギー事象を解析に導入するには、エネルギー閾値による単純な除去に代わり、複数の事象情報に基づくバックグラウンド抑制を行うことが不可欠である。

そこで本研究では、6MeV 以下の事象でも安定した再構成が可能な WIT の性能を活用し、6MeV のエネルギー閾値を適用せずに複数の事象パラメータを用いた選別によってバックグラウンドを除去することで、方向再構成に寄与する ES 事象数の増加を図ることを試みた。本研究では 6MeV より大きいエネルギーを持つ事象を高エネルギー事象、6MeV 以下の事象を低エネルギー事象と呼ぶ。

### 3.1 低エネルギー事象に対する選別

本節では、低エネルギー領域 (6MeV 以下) において ES 事象を保持しつつ検出器由来バックグラウンドを抑制するための選別条件を検討する。表 3.1 に各 SN 距離における中里モデルに基づく低エネルギー領域の SN ニュートリノ事象数の平均を示す。この表からも、6MeV 以下のエネルギー領域では他の SN ニュートリノ事象と比べて検出器由来バックグラウンド事象が支配的であることが分かる。そのため、ES 事象の寄与を維持したままこれらを除去する必要がある。

そこで本研究では、WIT で再構成された事象情報のうち、再構成の確からしさ Goodness、内部検出器壁面からの距離  $D_{\text{wall}}$ 、再構成エネルギー  $E$ 、および事象時刻  $t$  を用いて、適切な選別条件を検討した。複数の事象情報を用いるため、各パラメータについて有効と考えられる閾値候補を設定し、その組み合わせを評価することで最終的な選別条件を決定した。

表 3.1: 6MeV 以下のエネルギー領域での中里モデルに基づく超新星ニュートリノ反応事象数

相互作用名	事象数/割合 (4kpc)	事象数/割合 (10kpc)	事象数/割合 (20kpc)
BG	14990/87.45%	14990/97.85%	15171/99.47%
Neutron(IBD)	1119/6.53%	181/1.18%	44/0.29%
Positron(IBD)	503/2.94%	61/0.40%	14/0.10%
Electron(ES)	354/2.07%	58/0.38%	14/0.09%
Others	174/1.02%	28/0.18%	7/0.05%
合計	17141	15319	15252

### 3.1.1 再構成の確からしさ Goodness

2.2.3 節で述べたように、WIT では再構成された各事象に対して、再構成の確からしさを表す指標として Goodness を算出する。Goodness は 0 から 1 の範囲の値を取り、1 に近いほど再構成の信頼性が高いことを示す。図 3.1 に SN 距離 10kpc における低エネルギー事象の Goodness 分布を示す。

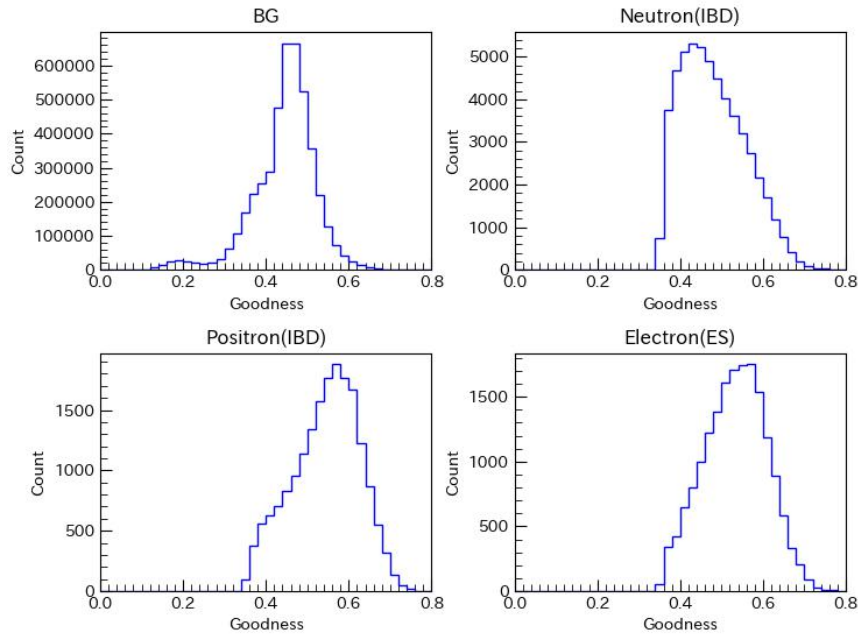


図 3.1: SN 距離 10kpc、300 個の MC データによる SN ニュートリノ事象毎に見た低エネルギー事象の Goodness 分布 (左上: バックグラウンド事象、右上:IBD 反応由来の中性子、左下:IBD 反応由来の陽電子、右下:ES 反応由来の電子)。バックグラウンド事象は 0.45 付近に、ES 事象は 0.5-0.6 にピークを示す。両者の分布は一部重なるため、閾値設定が結果に大きく影響する。

図 3.1 より、バックグラウンド事象は主に 0.3-0.6 の領域に分布し、特に 0.45 付近にピークを持つ。一方、ES 事象は 0.35-0.7 の範囲に広がり、0.5-0.6 に明確なピークを持つ分布を示す。ただし、ES 事象とバックグラウンド事象の分布は一部の領域で重なっており、閾値のわずかな違いで ES 事象の除去率も大きく変わるため、分布形状の観察のみから閾値を一意に決定することは難しい。

そこで、Goodness の閾値を定量的に決定するための評価指標として、本研究では Significance を導入した。これはバックグラウンド事象の統計揺らぎに対して信号事象がどの程度有意に抽出されているかを表す量であり、以下で定義する。

$$\text{significance} = \frac{N_{SG}}{\sqrt{N_{SG} + N_{BG}}} \quad (3.1)$$

ここで、 $N_{SG}$  は有効な信号事象数 (本研究では ES 事象数)、 $N_{BG}$  はバックグラウンド事象数である。図 3.2 に、Goodness 閾値を変化させた際の Significance 分布を示す。この分布から、いずれの SN 距離においても、Goodness 閾値が 0.55 付近で最大となることが分かる。

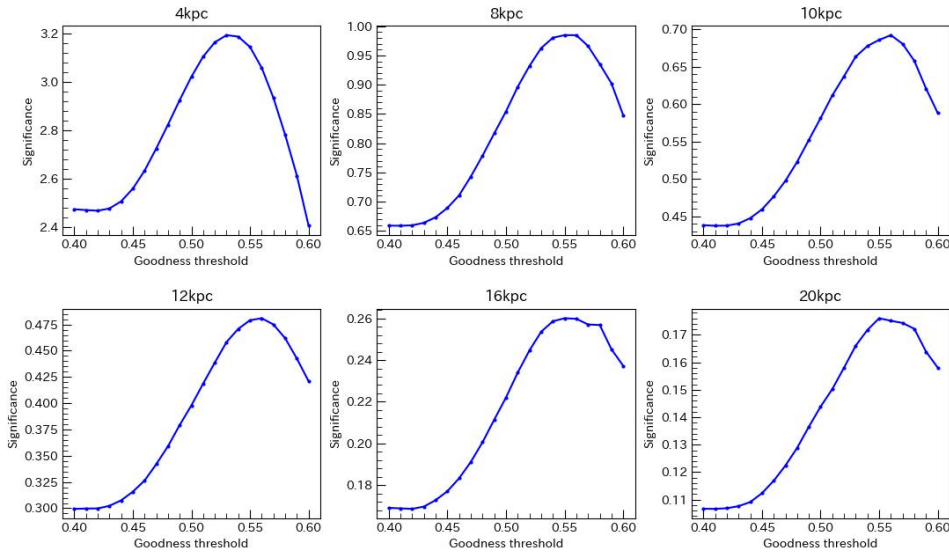


図 3.2: 低エネルギー事象に対して Goodness 閾値を変化させた際の Significance。いずれの SN 距離においても Goodness 閾値が 0.55 付近で最大となる。

さらに、図 3.3 に、ES 事象に対して SN 方向  $\hat{d}_{SN,true}$  と各事象の再構成運動方向  $\hat{u}_i$  のなす角の  $\cos$  値

$$\cos \theta_i = \hat{d}_{SN,true} \cdot \hat{u}_i \quad (3.2)$$

と Goodness の関係を表す 2 次元ヒストグラムを示す。この分布から、いずれの距離においても Goodness が 0.5-0.6 の領域に SN 方向と強い相関を持つ ES 事象が集中していることが分か

る。このため、Goodness 閾値を 0.55 に設定した場合、方向再構成に有効な ES 事象が過度に除去される可能性がある。以上を踏まえ、Goodness の閾値候補として 0.50 および 0.55 の二通りを採用した。

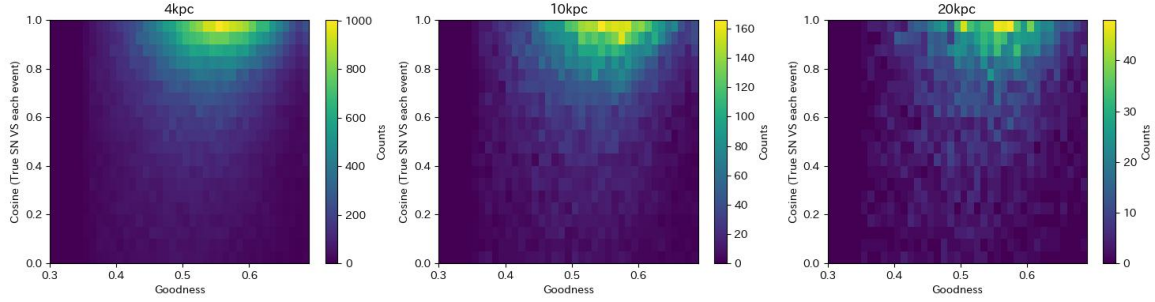


図 3.3: 低エネルギー事象における ES 事象の運動方向  $\hat{u}_i$  と SN 方向  $\hat{d}_{SN,true}$  の  $\cos \theta_i$  と再構成の確からしさ Goodness の 2 次元分布 (左:4kpc、中央:10kpc、右:20kpc)。いずれの SN 距離でも Goodness が 0.5-0.6 の領域に SN 方向と強い相関を持つ ES 事象が多く含まれる。

### 3.1.2 内部検出器壁面からの距離 $D_{wall}$

スーパーカミオカンデでは、外部検出器起源のバックグラウンド事象の混入や、検出器壁面近傍におけるチェレンコフ光の損失に伴う再構成精度の低下を避けるため、内部検出器 (ID) 壁面近傍で再構成された事象を除去することがある。本節では、低エネルギー事象に対する選別条件の一つとして、ID 壁面からの距離  $D_{wall}$  の閾値を検討する。

各事象の  $D_{wall}$  は、図 3.4 に示す ID 座標系に基づき以下の手順で算出した。

1. ID 半径  $R_{ID}$  と再構成 vertex 座標  $(x, y)$  から側面までの距離  $r_{xy}$  を次式で定義する。

$$r_{xy} = R_{ID} - \sqrt{x^2 + y^2} \quad (3.3)$$

2. ID の半高さ  $H_{ID}/2$  と vertex の  $z$  座標から、検出器の上下端までの距離  $r_z$  を

$$r_z = \frac{H_{ID}}{2} - |z| \quad (3.4)$$

と定義する。

3. 側面および上下端のうち最も近い壁面からの距離として、

$$D_{wall} = \min(r_{xy}, r_z) \quad (3.5)$$

を用いる。

図 3.5 に、SN 距離 10kpc における各 Goodness 閾値条件下での  $D_{wall}$  分布を示す。この分布から、Goodness による選別後に残存するバックグラウンド事象は特に ID 壁面から 300cm 以

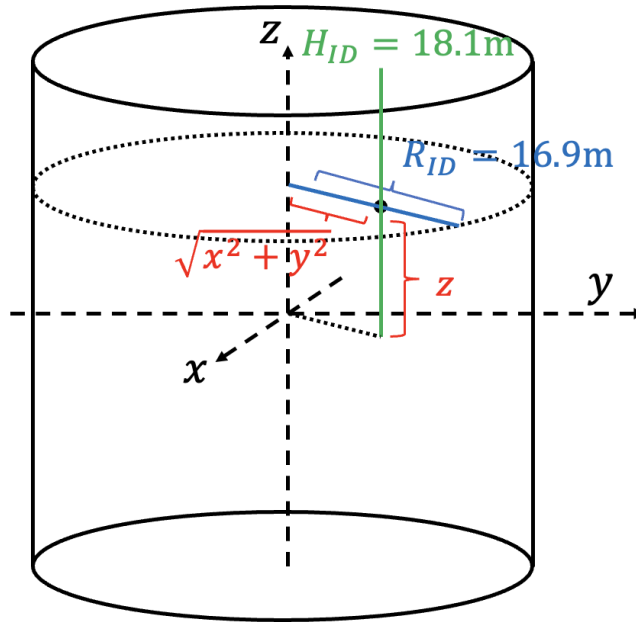


図 3.4: 内部検出器における座標の定義

内に集中しており、壁面近傍に複数のピーク構造が確認できる。これは、PMT や構造物に由来する外部放射性バックグラウンドによるものだと考えられる。

図 3.6 に、 $D_{wall}$  と各 ES 事象の  $\cos \theta_i$  の 2 次元ヒストグラムを示す。この分布から、いずれの SN 距離においても ES 事象は ID 内壁からの距離に依らず、SN 方向と強い相関を示すことが分かる。 $D_{wall}$  の閾値を 300cm に設定するとバックグラウンド事象を大幅に抑制できるが、同時に ES 事象も多く失う可能性がある。以上を踏まえ、 $D_{wall}$  の閾値候補として 200cm および 300cm の二通りを採用した。

### 3.1.3 再構成エネルギー $E$

本研究では、低エネルギー領域に多く含まれる ES 事象を有効に活用するため、現行の SNWATCH で導入されている 6MeV のエネルギー閾値に代えて、より低いエネルギー閾値を設定できるかを検討した。評価には、低エネルギー事象のエネルギー分布に加え、3.1.1 項で定義した各 ES 事象の  $\cos \theta_i$  分布を用いた。図 3.7 に、SN 距離 10kpc における低エネルギー事象のエネルギー分布を示す。この分布から、バックグラウンド事象のほぼ全てが 4MeV 以下の領域に集中していることが分かる。実際に、低エネルギー事象に含まれる全バックグラウンド事象に対する、 $E \leq 4\text{MeV}$  のバックグラウンド事象の割合は 99.2%であった。

次に、図 3.8 に、再構成エネルギーと各 ES 事象の  $\cos \theta_i$  の関係を示す。この分布から、3MeV 以下の領域と比較して 3-4MeV の領域において  $\cos \theta_i$  が高い、すなわち真の SN 方向と強い相関を持つ事象が多く分布していることが分かる。実際に、3MeV 以下の ES 事象は低エ

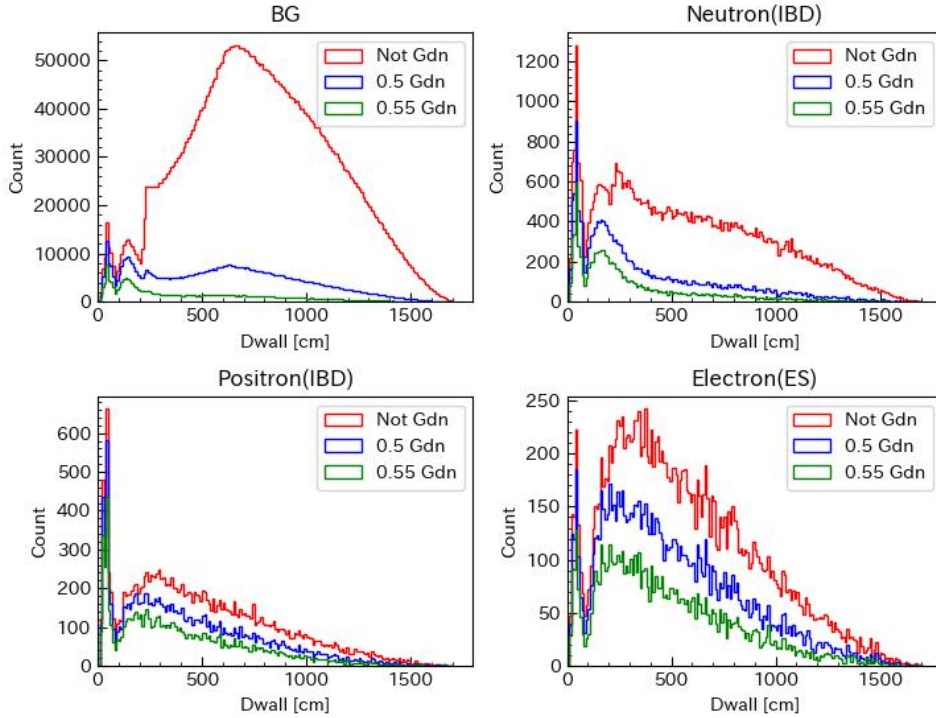


図 3.5: SN 距離 10kpc、300 個の MC データによる各 Goodness 閾値条件下での低エネルギー事象の Dwall 分布。Goodness 閾値条件下では、SN ニュートリノ事象と比べて、多数のバックグラウンド事象が ID 壁面近傍に分布している。

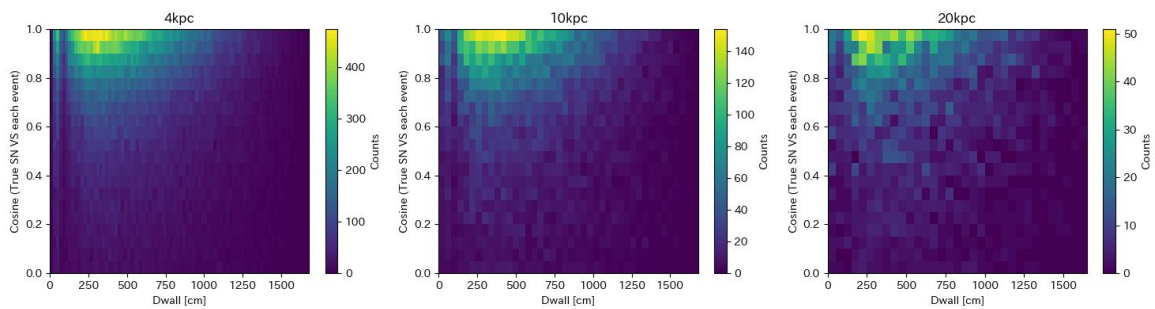


図 3.6: 低エネルギー事象における ES 事象の運動方向  $\hat{u}_i$  と SN 方向  $\hat{d}_{SN,true}$  の  $\cos \theta_i$  と ID 壁面からの距離 Dwall の 2 次元分布 (左:4kpc、中央:10kpc、右:20kpc)。いずれの SN 距離においても Dwall が 200cm から 300cm の範囲に多くの SN 方向と強い相関のある ES 事象が分布している。

エネルギー領域に含まれる ES 事象の 14%程度であるのに対し、3-4MeV の ES 事象は 28%程度を占めており、方向推定に有効な ES 事象が相対的に多く含まれる。以上より、バックグラウンドの大部分を抑制しつつ、方向推定に寄与する ES 事象の除去を最小限に抑えるため、本研究では低エネルギー事象に対するエネルギー閾値を 3MeV に設定した。

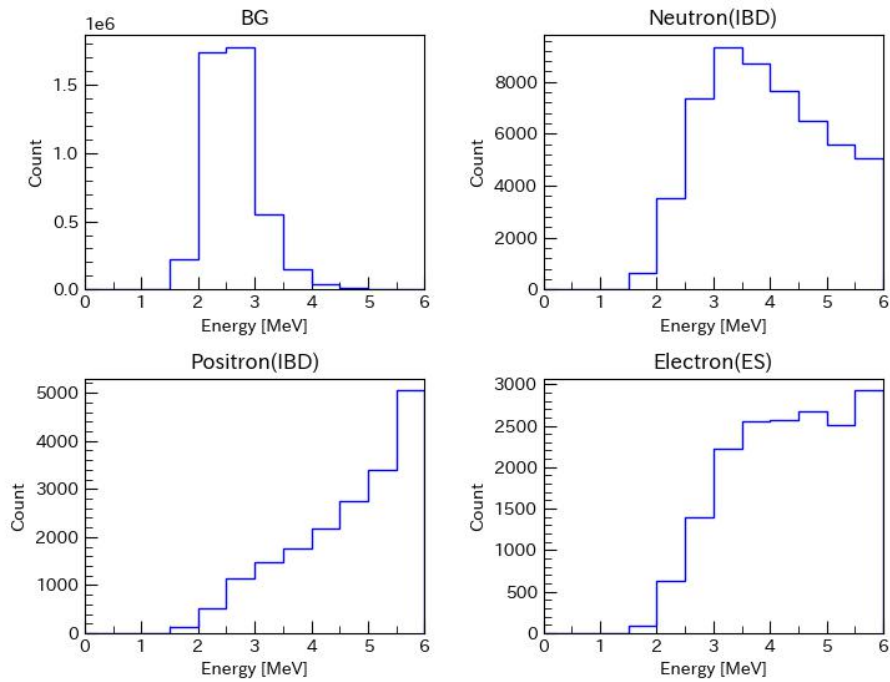


図 3.7: SN 距離 10kpc、300 個の MC データにおける低エネルギー事象のエネルギー分布。ほぼ全てのバックグラウンドが 4MeV 以下の領域に分布している。

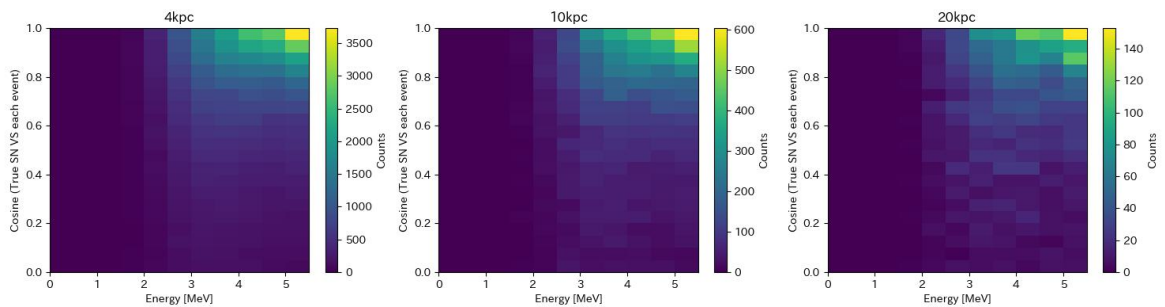


図 3.8: 低エネルギー事象における ES 事象の運動方向  $\hat{u}_i$  と SN 方向  $\hat{d}_{SN,true}$  の  $\cos \theta_i$  と再構成エネルギー  $E$  の 2 次元分布 (左:4kpc、中央:10kpc、右:20kpc)。いずれの SN 距離においても 3MeV 以上の領域に SN 方向と強い相関を持つ ES 事象が多く分布する。

### 3.1.4 再構成事象時刻 $t$

最後に、事象の再構成時刻情報  $t$  を用いたバックグラウンド抑制を検討した。SN ニュートリノ事象はバーストとして短時間に集中して観測される一方で、検出器由来バックグラウンドは超新星爆発と無関係であるため、時間範囲全体に渡ってほぼ一様に分布すると期待される。

図 3.9 に、SN 距離 10kpc における各 Goodness 閾値条件下での再構成時間分布を示す。この分布から、バックグラウンド事象は時間範囲全体に渡ってほぼ一様に分布する一方で、SN ニュートリノ事象は概ね 1 秒以内に集中することが分かる。従って、一定時間以内の事象に限定することで、バックグラウンド事象を抑制し、ES 事象を効率的に抽出できる可能性がある。

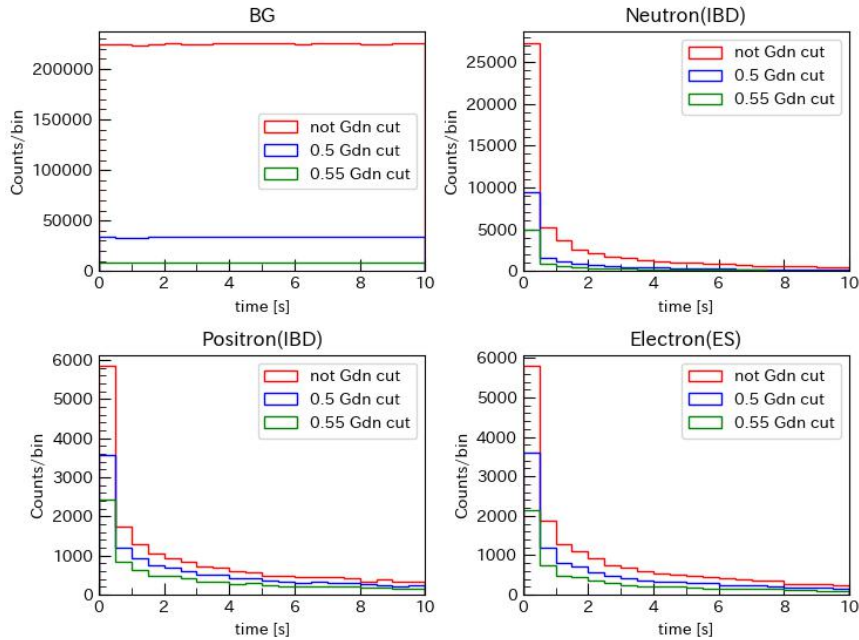


図 3.9: SN 距離 10kpc、MC データ 300 個による Goodness 閾値下での低エネルギー事象の時間分布。SN ニュートリノ事象は 1 秒以内にピークを示し、バックグラウンドは一様に分布する

ただし、時間分布のみから閾値を一意に決定することは困難であるため、Goodness、 $D_{\text{wall}}$  およびエネルギーの候補閾値を組み合わせた条件下で、時間閾値を変化させた際の Significance(式 3.1) を評価した。図 3.10 にその結果を示す。図 3.10 から、いずれの SN 距離においても Goodness 閾値を 0.50、 $D_{\text{wall}}$  閾値を 200cm とした場合に Significance が高くなる傾向が確認できた。また、SN 距離が大きくなるにつれ、Significance が最大となる時間閾値は小さくなる傾向も見られた。近距離 SN では観測事象数が多く、ES 事象数も十分確保される一方、遠距離 SN では ES 事象数が少なく、閾値設定の影響も受けやすい。そのため本研究で

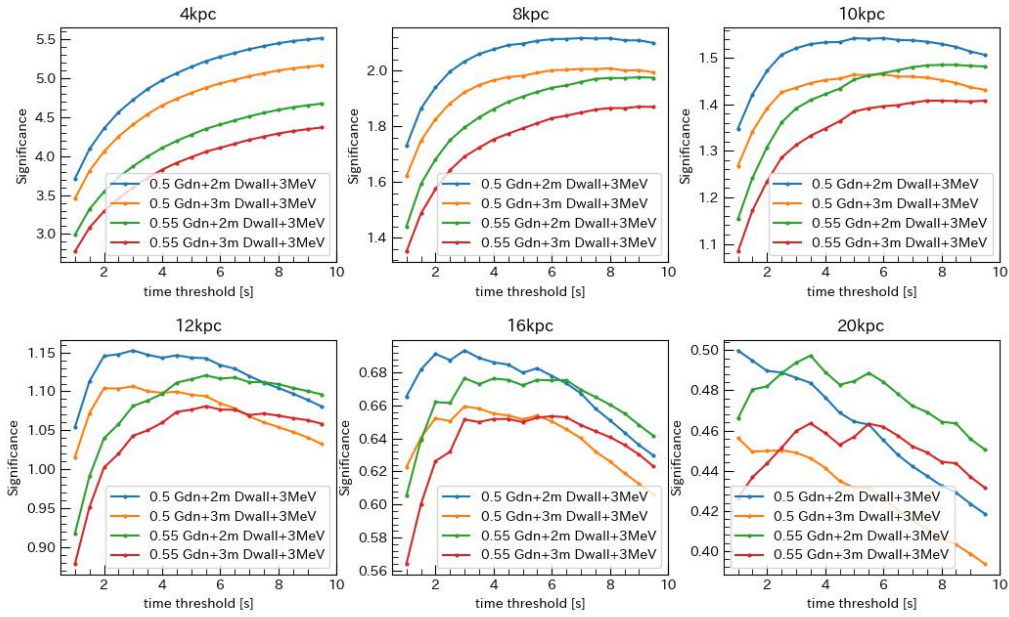


図 3.10: 低エネルギー事象に対して各パラメータの候補閾値を組み合わせ求めて Significance。いずれの SN 距離においても Goodness 閾値が 0.5、Dwall 閾値が 2m の時に Significance が高くなる。Significance が最大となる時間閾値は SN 距離によって異なる。

は、遠距離 SN における Significance の振る舞いを重視し、10-20kpc の範囲で Significance が安定して高い値を示した  $t = 3$  秒を時間閾値として採用した。

以上を踏まえ、本研究では低エネルギー領域に対して、事象再構成の確からしさ  $\text{Goodness} \geq 0.5$ 、ID 壁面からの距離  $\text{Dwall} \geq 200\text{cm}$ 、再構成エネルギー  $E \geq 3$ 、事象時刻  $t \leq 3\text{s}$  となる事象を選別した。

### 3.2 高エネルギー事象に対する選別

表 3.2 に、各超新星距離での中里モデルに基づく高エネルギー領域 ( $E > 6\text{MeV}$ ) の SN ニュートリノ事象数の平均を示す。高エネルギー領域では、低エネルギー領域と比較して検出器由来バックグラウンドの寄与が十分に小さいことが分かる。しかし、高エネルギー領域にも SN ニュートリノ到来方向と相関を持たない事象が含まれるため、方向再構成に有効な ES 事象の寄与を高める目的で、閾値設定の検討を行った。

高エネルギー領域に対して、低エネルギー事象と同様に多数の閾値を設定すると、かえって方向再構成に有効な ES 事象を過度に除去する可能性がある。そこで高エネルギー事象については、事象再構成の信頼性に直結する Goodness と Dwall を用いて事象選別をする方針とした。

表 3.2: 6MeV より高いエネルギー領域での中里モデルに基づく超新星ニュートリノ事象数

相互作用名	事象数/割合 (4kpc)	事象数/割合 (10kpc)	事象数/割合 (20kpc)
BG	4/0.05%	4/0.26%	4/0.84%
Neutron(IBD)	232/2.16%	32/1.73%	7/1.62%
Positron(IBD)	8900/82.70%	1541/82.65%	394/82.40%
Electron(ES)	713/6.63%	119/6.38%	29/6.24%
Others	912/8.47%	167/8.98%	42/8.89%
合計	17141	15319	15252

まず、高エネルギー事象における Goodness による事象選別の有効性を確認する。図 3.11 に、代表例として SN 距離 10kpc における高エネルギー事象の Goodness 分布を示す。この分布から、バックグラウンド事象および IBD 事象由来の中性子は主に 0.3-0.8 の広い範囲に分布し、特にバックグラウンド事象は 0.45 付近に集中していることが分かる。ただし、それらの事象数は低エネルギー事象と比較して明らかに少ない。一方、ES 事象は主に 0.5-0.8 の範囲に分布しており、Goodness が大きい領域により多く集中する傾向が確認できる。以上より、高エネルギー事象においても Goodness に基づく事象選別は有効であると判断できる。次に、

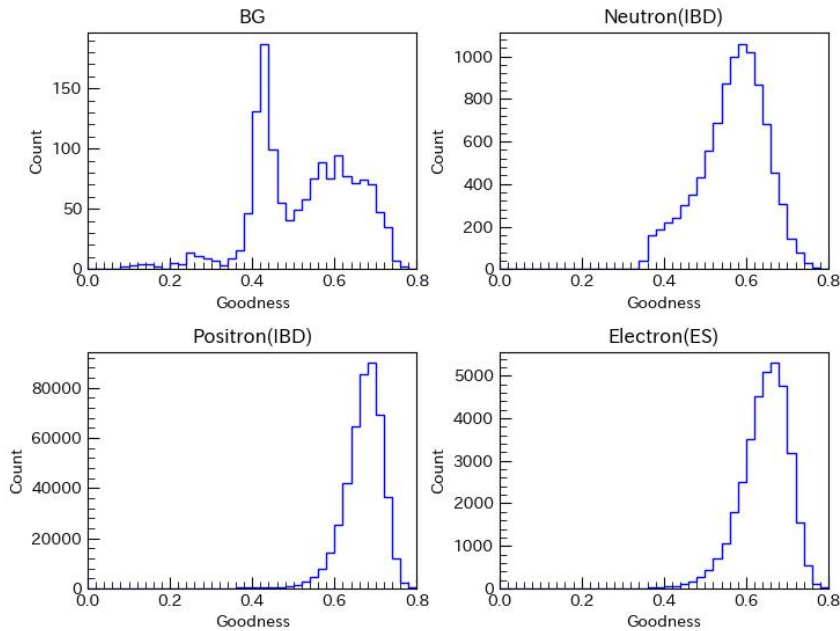


図 3.11: SN 距離 10kpc、300 個の MC データによる高エネルギー事象の Goodness 分布。ES 事象は Goodness>0.5 に分布する一方、バックグラウンド事象と IBD 由来中性子事象は Goodness<0.5 の領域にも分布する。このため、Goodness 選別が有効と示唆される。

高エネルギー事象における  $D_{\text{wall}}$  の必要性を検討する。図 3.12 に、SN 距離 10kpc における各 Goodness 閾値条件下での  $D_{\text{wall}}$  分布を示す。この分布から、ID 壁面近傍においても、SN ニュートリノ反応由来事象が多数存在する一方で、バックグラウンド事象数は非常に少ないことが分かる。また、バックグラウンド事象の  $D_{\text{wall}}$  分布では、Goodness の閾値を設定しない場合に  $D_{\text{wall}}=1400\text{cm}$  付近に顕著なピークが現れる。このピークの起源は現時点では不明であるが、Goodness 閾値を 0.5 に設定することでこの構造は除去されることから、Goodness による事象選別が有効であることが確認できる。さらに、 $D_{\text{wall}}$  による ES 事象の損失の程度を

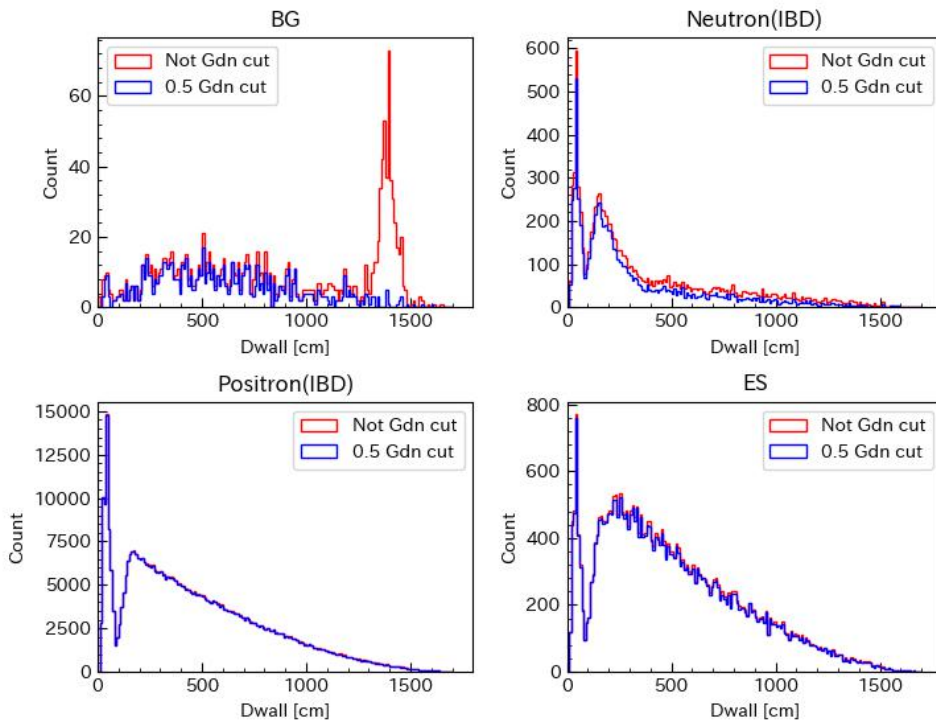


図 3.12: SN 距離 10kpc、300 個の MC データによる Goodness 閾値条件下での高エネルギー事象の  $D_{\text{wall}}$  分布。低エネルギー事象と比べてバックグラウンドが少ない。

確認するために、図 3.13 に、 $D_{\text{wall}}$  と各 ES 事象の  $\cos \theta_i$  の関係を表す二次元ヒストグラムを示す。この分布から、ID 壁面近傍に位置する ES 事象においても、SN 方向と強い相関を持つことが分かる。したがって、 $D_{\text{wall}}$  による選別は方向再構成におけるバックグラウンド事象の抑制には寄与する一方で、有効な ES 事象も同時に除去する可能性がある。そこで、 $D_{\text{wall}}$  の閾値を 0cm および 200cm の二通りとして、横軸を Goodness 閾値、縦軸を Significance(式 3.1) とした分布を作成した。図 3.14 にその結果を示す。図 3.14 より、いずれの SN 距離においても、 $D_{\text{wall}}$  閾値を設定しない場合の方が、設定した場合と比較して高い Significance を示すこ

とが分かった。これは、高エネルギー領域では壁面近傍のバックグラウンド寄与が小さい一方で、Dwall 閾値によって ES 事象の損失が生じるため、結果として信号抽出性能 ( $S/\sqrt{S+B}$ ) が低下することに起因すると考えられる。また、Goodness 閾値を 0.3 から 0.5 の範囲で変化させても、Significance はほぼ同程度の値を示している。

以上を踏まえ、本研究では高エネルギー事象に対して Goodness  $\geq 0.5$  を満たす事象を選別し、Dwall 閾値は設定しない方針を採用した。

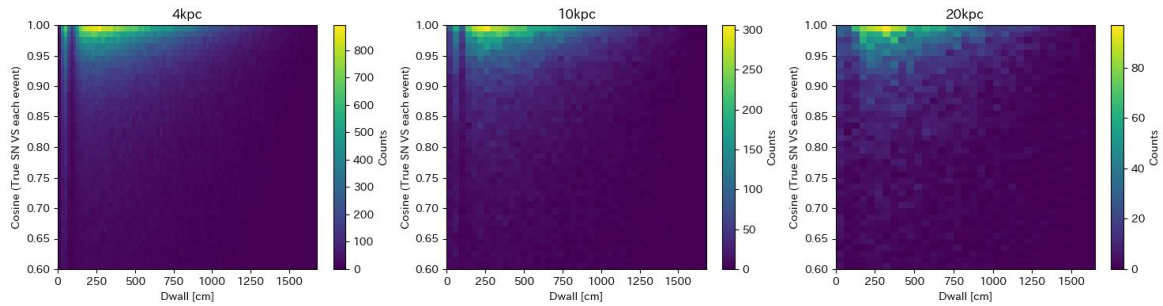


図 3.13: 高エネルギー事象における ES 事象の運動方向  $\hat{u}_i$  と SN 方向  $\hat{d}_{SN,true}$  の  $\cos \theta_i$  と ID 壁面からの距離 Dwall の 2 次元分布 (左:4kpc、中央:10kpc、右:20kpc)。いずれの SN 距離においても ID 壁面近傍でも SN 方向と強い相関を持つ ES 事象が多く分布する。

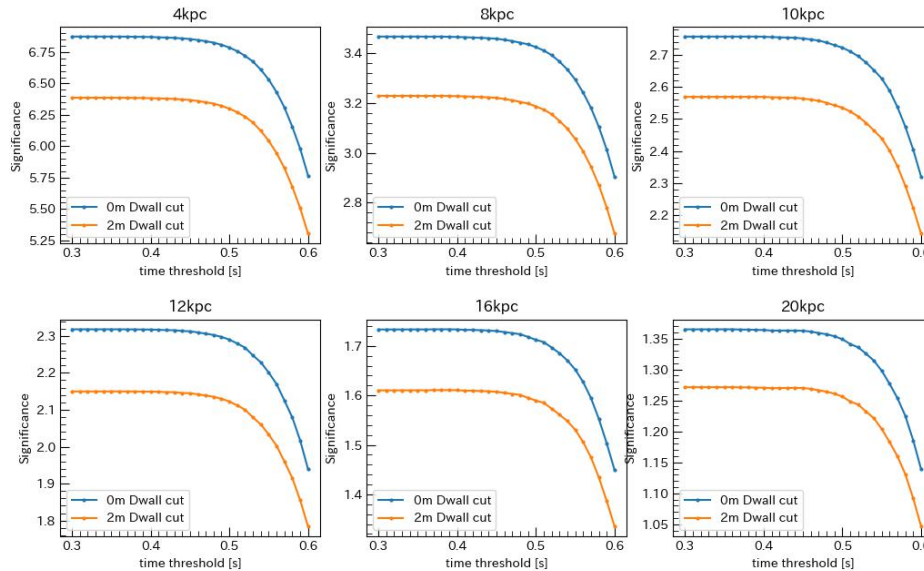


図 3.14: 高エネルギー事象に対する Dwall 候補閾値を組み合わせ求めて Significance。いずれの SN 距離においても Dwall 閾値を設定しない方が Significance が高くなる。また、Goodness 閾値 0.5 以下で安定して高い Significance を示す。

## 第 4 章 DBSCAN fitter

本章では、本研究で超新星爆発の方向再構成のために導入した密度に基づくクラスタリング手法 DBSCAN と、それを用いた方向再構成手法 (DBSCAN-fitter) について述べる。現行の SNWATCH における方向再構成手法が事象分布全体を用いた統計的推定に基づいているのに対し、本研究では、電子散乱 (ES) 事象が形成する局所的な高密度領域に着目し、その空間構造を直接抽出することを目的としている。本章では、DBSCAN-fitter の各アルゴリズムについて概説するとともに、各種パラメータの最適化については 5 章で述べる。

### 4.1 DBSCAN

機械学習とは、データから規則性や特徴を自動的に抽出し、未知のデータに対する予測や分類を行う手法である [16]。機械学習は一般に、教師あり学習、教師なし学習、半教師あり学習に分類される。

超新星ニュートリノの到来方向と強い相関を持つ ES 事象が、再構成された事象の運動方向分布において、SN 方向の周辺に局所的な高密度領域を形成する点に着目し、密度に基づくクラスタリング手法 DBSCAN (Density-Based Spatial Clustering of Applications with Noise) を採用した。DBSCAN は教師なし学習の一種であり、事象のラベル情報や分布形状を仮定することなく高密度領域を直接クラスタとして抽出できる特徴を持つ。SN ニュートリノバースト観測において、各事象がどのニュートリノ反応から生成されたかを判別することは困難であるため、ラベル情報を必要としない教師なし学習は本研究の解析方針と整合的である。DBSCAN では、高密度領域に属するデータポイントをコアサンプルと呼び、クラスタ探索には、近傍の探索半径を定めるパラメータ  $\epsilon$  と、コアサンプルと判定するために必要な最小データ数  $\text{min\_samples}$  の二つを与える必要がある。以下では、図 4.1 に示した DBSCAN のアルゴリズムに基づいて、その手順を説明する。

1. 特徴量空間上 (本研究では事象の運動方向分布) にデータセットを配置し、まだクラスタに割り当てられていない任意のデータポイントの一つを選択する。(図 4.1 の左上)
2. 選択したデータポイントを中心として、半径  $\epsilon$  以内に存在する近傍点を探し、その数を数える。(図 4.1 の中央上)
3. 近傍点の数が  $\text{min\_samples}$  未満である場合、そのデータポイントはノイズ候補として扱われる。一方、 $\text{min\_samples}$  以上である場合、そのデータポイントはコアサンプルと判定

され、新たなクラスタが生成される。(図 4.1 の右上)

4. コアサンプルの  $\text{eps}$  近傍に存在するデータポイントを同じクラスタに割り当てる。これらの点に対し 2 と 3 の操作を行い、その中でコアサンプルが含まれていた場合、その点を中心としてさらに近傍探索を行い、クラスタを再帰的に拡張する。(図 4.1 の左下)
5. クラスタの拡張が完了した後、まだ処理されていないデータポイントを選択し、同様の手順を全てのデータポイントに対して繰り返すことで、最終的な分類を終了する。(図 4.1 の右下)

DBSCAN は局所的な密度構造を直接抽出できる性能を持つため、遠距離 SN のように事象が少ない状況下でも安定したクラスタ抽出が可能であると期待される。一方で、局所的な事象の偏りが最終的な方向推定に影響を与える可能性もある。

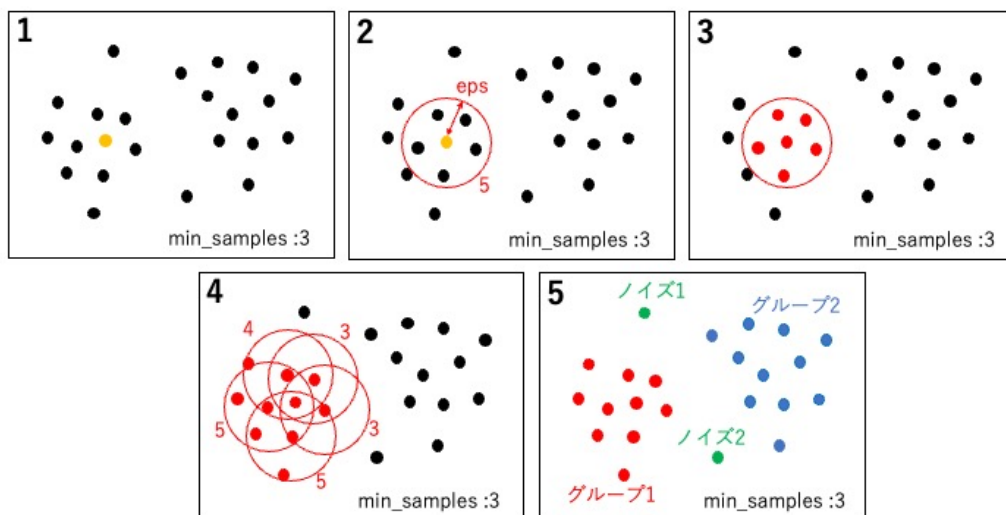


図 4.1: DBSCAN のアルゴリズム (ここでは  $\text{min\_samples}$  を 3 にした状況を考えている)

## 4.2 解析フロー

DBSCAN-fitter は WIT によって再構成された事象に対し、第 3 章で設定した各種事象情報に基づく選別、事象数に基づく DBSCAN パラメータ 9 組の推定、DBSCAN の適用によるクラスタの抽出、各パラメータセット毎に方向を再構成、Estimator による最終結果の選別、という過程を通して、超新星爆発の方向を再構成する。本節では、設定した各アルゴリズムについて概説する。

### 4.2.1 設定した閾値による事象の選別

SK で SN ニュートリノバーストが検出された場合、まずガドリニウムを用いた IBD 事象タグ付けにより、中性子を伴う事象を除去する。次に、再構成された全事象に対して、第 3 章で述べた各種事象情報に基づく閾値を適用し、検出器由来バックグラウンド事象を除去する。これらの選別操作を適用した後に得られる全事象の再構成運動方向分布を、図 4.2 に示す。

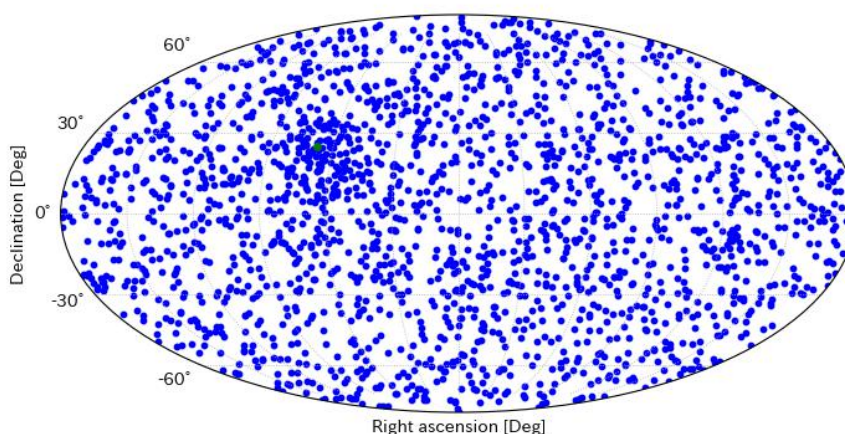


図 4.2: IBD タグ付けと各種事象情報に基づく選別後の再構成運動方向分布

### 4.2.2 パラメータ推定

パラメータ推定式 (表 5.3 と表 5.4) に基づき、選別後の事象数を入力として、DBSCAN パラメータである `eps` および `min_samples` を算出する。これらの推定式は、SN ニュートリノ事象に関するモンテカルロシミュレーションによって生成した訓練データを用いて、DBSCAN パラメータと角度分解能との関係から導出したものである。その詳細については、5.3 節で述べる。

DBSCAN 解析において理想的には、訓練データを用いて SN 方向を最も高精度に再構成できるパラメータと、SN ニュートリノ事象に由来する特徴量との間に強い相関を見出し、それに基づいてパラメータを一意に決定することが望ましい。しかし現時点では、最良の角度分解能を与えるパラメータと強く相関する単一の特徴量、あるいは複数の特徴量の組み合わせを見出すには至っていない。また、SN ニュートリノバースト観測においては、SN までの距離や検出器由来バックグラウンドの影響により、再構成される事象数や ES 事象の密度分布が SN 事象ごとに大きく変化する。そのため、単一のパラメータセットを用いた解析では、ある条件下では有効であっても、他の条件下ではクラスタ抽出に失敗する可能性がある。このような事象状況の多様性を考慮し、本研究では複数の DBSCAN パラメータセットによる探索的解析を採

用した。この手法により、SN 方向を高精度に再構成できるパラメータセットが含まれる確率を高めるとともに、SN ニュートリノ事象に由来するクラスタをより安定的に抽出することを狙っている。本研究では解析手法の設計段階において、パラメータ数増加による計算時間の増大と失敗率低減とのトレードオフを考慮し、9 組のパラメータセットを用いて解析を行う方針とした。9 組のパラメータセットは、3 種類の eps(小さい順に Bottom\_eps、Middle\_eps、Up\_eps と定義する) と、各 eps に対して設定した 3 種類の min\_samples の組み合わせから構成されている。SN までの距離 10kpc の場合に推定されたパラメータセットの一例を表 4.1 に示す。

表 4.1: SN 距離 10kpc の SN 事象 (事象数:1992) に対し、全エネルギー事象を対象とした際に表 5.4 を基に推定されるパラメータ表

eps	(推定 min_samples) $\times$ 0.925	推定 min_samples	(推定 min_samples) $\times$ 1.075
(Up) 0.26	61	66	70
(Middle)0.38	113	123	132
(Bottom)0.52	182	197	211

### 4.2.3 DBSCAN によるクラスタ選別

パラメータ推定によって算出される eps と min\_samples の組み合わせを用いて、全事象の再構成運動方向分布に対して DBSCAN を適用する。この操作によって、各パラメータセット毎に ES 事象が集中したと想定されるクラスタと、それ以外の領域に分類される。その結果を可視化した分布を図 4.3 に示す。DBSCAN によって各事象に付与されたクラスタラベルは、次項で述べる SN 方向再構成で用いられる。なお、本解析システムでは、DBSCAN の適用対象となる再構成運動方向を全て単位ベクトルとして扱い、方向ベクトル間のユークリッド距離を距離指標として用いる。単位ベクトル間のユークリッド距離は方向の角距離と単調な関係を持ち、計算コストも小さいため、本研究ではこれを採用した。

### 4.2.4 SN 方向再構成

各パラメータセット毎に、DBSCAN による分類結果に基づいて方向再構成を行う。方向再構成手法として最も導入が簡単なのは、抽出されたクラスタ事象の運動方向を単純平均する手法である。しかし、抽出されたクラスタは必ずしも真の SN 方向に対して対象的な分布を持つとは限らないため、単純平均では方向再構成精度が低下する可能性がある。そこで本研究では、真の SN 方向近傍に集中する事象ほど局所密度が高いことを仮定し、教師あり学習手法の

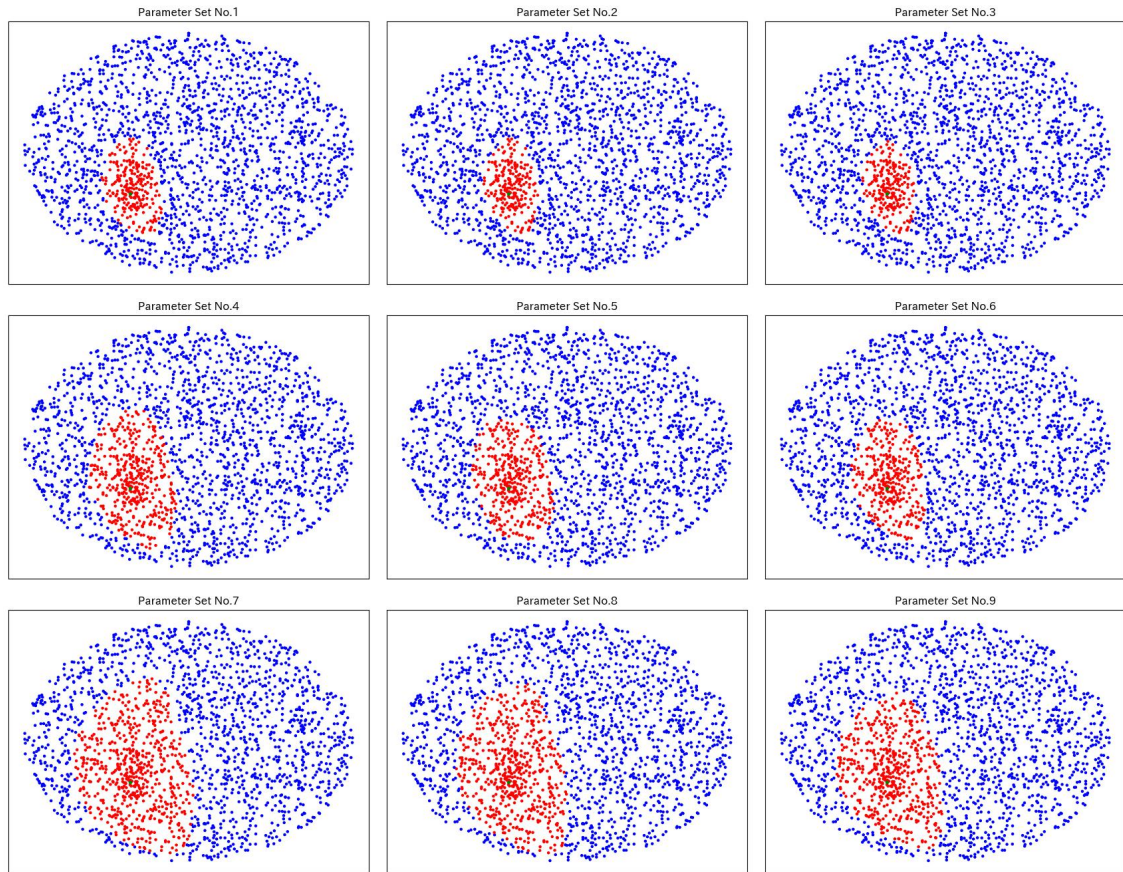


図 4.3: DBSCAN による選別後の全事象運動方向分布 (赤:DBSCAN が抽出したクラスタ領域)

一つである k-Nearest Neighbors(kNN) を用いた距離に基づく評価方法を採用し、方向再構成精度の向上を図った。

kNN はデータ点間の距離情報を利用し、各データポイント近傍の局所的な分布特性を評価する手法がある。本研究ではこの性質を用いて以下の手順により、運動方向分布上における各事象の局所密度を評価するための距離を算出した。図 4.4 にその模式図を示す。

1. 再構成された全データポイントに対して、各データポイントと他の全データポイントとのユークリッド距離を計算する。
2. 各データポイントについて、距離が近い順に指定した数の近傍点を抽出する。この近傍点の数を  $n\_neighbors$  と定義する。
3. 各データポイントに対して抽出された  $n\_neighbors$  個の近傍距離の平均値を計算する。本

研究では、この平均距離を MeanKdistance と呼び、当該データポイント周辺の局所密度を表す指標として用いる。なお、平均を取る操作は、本研究独自の処理であり、標準的な k-Nearest Neighbors アルゴリズムには含まれていない。

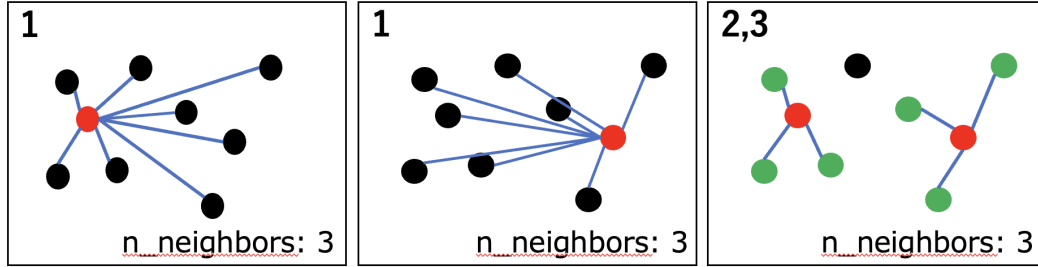


図 4.4: k-Nearest Neighbors のアルゴリズム (赤点是对象のデータポイント、緑は対象のデータポイントに最も近い 3 個の近傍点、青はデータポイント間の距離を示す)

図 4.5 は、SN までの距離 10kpc における各事象の MeanKdistance と、事象の運動方向と SN 方向の  $\cos \theta_i$  との関係を表した散布図である。この分布から、MeanKdistance が小さい事象ほど SN 方向と強い相関を持つことが分かる。

この傾向を方向再構成に反映させるために、本研究では MeanKdistance の各ビンにおいて、SN 方向との  $\cos$  値が所定の閾値以上となる事象の割合を表す指標として Fraction を定義した。この Fraction は、当該 MeanKdistance 範囲に属する事象が SN 方向と相関を持つ度合いを表す量だと見なせる。さらに、この Fraction を基づき、MeanKdistance に依存した重み付け関数  $F_i(\text{MeanKdistance})$  を定義した。

$$F_i(\text{MeanKdistance}) = \frac{1}{1 + \exp(a \cdot (\text{MeanKdistance} - b))} \quad (4.1)$$

最終的に、9 通りの解析結果それぞれに対し、DBSCAN によって付与されたクラスタラベル  $L_i$  (クラスタに属する事象に対して  $L_i = 1$ , それ以外の事象に対して  $L_i = 0$ ) と重み付け関数  $F_i(\text{MeanKdistance})$  を用いて、以下の式 4.2 に従って SN 方向を再構成する。

$$\hat{d}_{\text{SN}}^{\text{recon}} = \frac{\sum_i F_i \cdot L_i \cdot \hat{u}_i}{|\sum_i F_i \cdot L_i \cdot \hat{u}_i|} \quad (4.2)$$

ここで  $\hat{u}_i$  は、各事象の再構成された運動方向ベクトルを示す。本再構成手法では、いくつかのパラメータを設定する必要があるが、それらの定義および最適化については第 5 章で詳しく述べる。

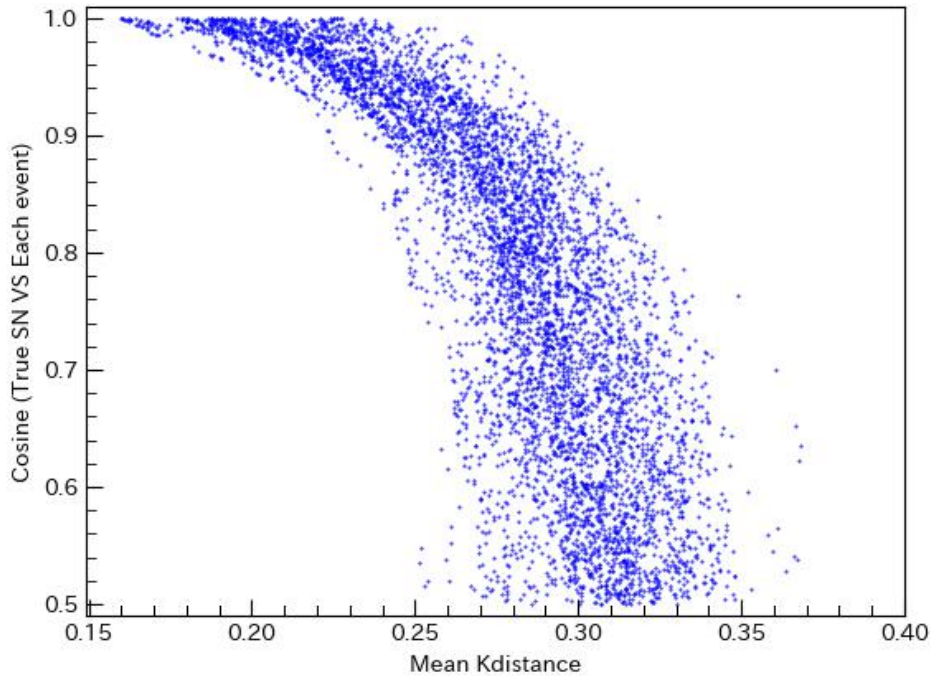


図 4.5: 全事象の運動方向  $\hat{u}_i$  と SN 方向  $\hat{d}_{SN,true}$  の  $\cos \theta_i$  と MeanKdistance による散布図。MeanKdistance が小さい事象ほど、SN 方向と強い相関を持つことが分かる。

#### 4.2.5 Estimator による結果選別

以上の手順をある SN 事象の運動方向分布に対して適用することにより、9 通りのクラスタレベルと、9 通りの再構成された SN 方向が得られる。これらの結果の中から、最も確からしく SN 方向を示す解析結果を選別するために、Minimum Distance Estimator(MDE) と呼ぶ独自の評価指標を用いた。MDE は SN 方向そのものを新たに推定するものではなく、複数の再構成結果の中から最も妥当な結果を選択するための結果選別用 Estimator である。

4.2.4 項で述べたように、再構成された運動方向分布上では、ES 事象の特性により真の SN 方向近傍で集中的な分布が期待される。したがって、SN 方向を高精度に再構成できた場合、その方向近傍には高い事象密度が形成されると考えられる。MDE はこの特徴に着目し、再構成した方向周辺に存在する事象を用いて、局所的な事象密度を評価することで、最も密度が高くなる結果を最終的な SN 方向として抽出する手法である。以下に、本研究で用いた MDE のアルゴリズムを示す。

1. 運動方向分布に対して、一つの再構成された SN 方向を基準とし、その方向ベクトルとの

ユークリッド距離が近い順に、予め設定された数 ( $N_{\text{MDE}}$ ) の事象を抽出する。(MDE を適用する際に用いる事象数の最適化については第 5 章で述べる。)

2. 再構成された SN 方向ベクトルと、抽出した各事象の再構成運動方向ベクトルとのユークリッド距離をそれぞれ算出し、その総和を求める。
3. 他の再構成方向に対しても同様の操作を行い、得られた 9 通りの距離和を比較する。その中で距離和が最も小さい結果を、最も確からしい SN 方向を与える結果として選別する。

# 第 5 章 評価指標と DBSCAN-fitter における パラメータの最適化

本章では、第 6 章以降で用いる DBSCAN-fitter の性能評価指標と、前章で述べた DBSCAN-fitter による超新星方向再構成手法において、実際に解析に用いる各種パラメータ最適化について述べる。SN ニュートリノバースト観測では、SN までの距離やバックグラウンド条件に応じて観測される事象数や事象分布が大きく変化するため、固定的なパラメータ設定では安定した方向再構成を行うことが困難である。そこで本研究では、モンテカルロシミュレーションによって生成した訓練データを用い、観測される事象数に基づいて DBSCAN、k-Nearest Neighbors、および Minimum Distance Estimator に関する各種パラメータを推定する手法を構築した。

本章では、SN ニュートリノモンテカルロシミュレーション、性能評価に用いる評価指標、各アルゴリズムのパラメータ最適化という順で説明を行う。ここで定義されたパラメータ推定式および評価指標は、次章において方向再構成性能の評価に用いられる。

## 5.1 モンテカルロシミュレーション (MC)

人類がこれまでに観測した超新星ニュートリノバーストは、1987 年にカミオカンデで観測された SN1987A のみである。そのため、超新星方向再構成手法の性能評価はモンテカルロシミュレーションに基づいて行う必要がある。本研究では、SN モデルに従う超新星爆発に伴うニュートリノフラックスの生成から、SK におけるニュートリノと標的原子核の反応、検出器の応答、PMT のデータを用いた事象の再構成までを再現したシミュレーションデータを用いた。

SN ニュートリノフラックスは、SN モデル、超新星までの距離、天球上における SN 方向、およびニュートリノ振動に関するパラメータを入力することによって生成される [14]。SN モデルの選択により、ニュートリノフレーバー、エネルギー、発生時間の情報をもつニュートリノフラックスが与えられる。

SN におけるニュートリノと標的原子核の相互作用生成には、SKSNSim(Super-Kamiokande Supernova Simulation) が用いられる。SKSNSim では、ニュートリノフラックスに基づき、SN 検出器内で期待される水標的との相互作用が生成される。生成された各ニュートリノ反応事象には、反応発生時刻、発生位置、ニュートリノエネルギー、相互作用の種類、および生成粒子の運動方向やエネルギーなどの真値情報が付与される。

SKSNSim で生成されたニュートリノ反応事象は、SKG4(Super-Kamiokande GEANT4-based

simulation) によって処理され、SK 検出器における応答が再現される。SKG4 では、チェレンコフ光の生成および伝播、PMT による光子検出、ならびにその応答がシミュレーションされる。また、ガドリニウム濃度 0.033%における中性子捕獲過程も実装されており、IBD 反応に伴う中性子捕獲事象が再現される。

さらに、実データに基づく放射性バックグラウンドおよび PMT の暗電流などのノイズを付加することで、より現実的な検出環境を再現する。これらのデータに対して、IBD タグ付け、および WIT による事象の反応発生位置 (vertex)、エネルギー、最初のニュートリノ事象からの時刻、運動方向などの再構成が行われ、最終的なモンテカルロシミュレーションデータが得られる。

本研究では、Gd 濃度 0.033%の状況下で、中里モデルに基づく超新星ニュートリノのモンテカルロシミュレーションを用いた。その作成にあたっては、前駆星質量、ショック再活性化時間および金属量を設定し、それぞれ  $M = 20M_{\odot}$ ,  $100\text{ms}$ ,  $Z = 0.02$  とした。これらの値は、先行研究において広く用いられている設定であり、特に前駆星質量  $M = 20, M_{\odot}$  は、SN1987A の起源星に対応する現実的なモデルである。また、ニュートリノ振動については、現在最も有力とされている正階層 (Normal Mass Ordering; NMO) を仮定した。また、本研究で作成した MC データにおいて、Gd による IBD 事象のタグ付け率は 55%である。加えて事象選別については、第 3 章で述べた条件を適用した。再構成エネルギーが 6MeV を超える事象に対しては、再構成の確からしさを示す  $\text{Goodness} \geq 0.5$  の事象を選別した。一方、6MeV 以下のエネルギー領域では、 $\text{Goodness} \geq 0.5$ 、内部検出器壁面からの距離  $D_{\text{wall}} \geq 2\text{m}$ 、再構成エネルギー  $E \geq 3\text{MeV}$ 、事象発生時刻  $t \leq 3\text{s}$  となる事象を選別している。

## 5.2 評価指標

### 5.2.1 角度分解能

DBSCAN-fitter による SN 方向再構成精度は、式 5.1 で定義される真の SN 方向と再構成された方向との角度差  $\Delta\theta_{SN}$  を用いて評価した。ここで、真の SN 方向を  $\hat{d}_{SN}^{\text{true}}$ 、再構成された方向を  $\hat{d}_{SN}^{\text{recon}}$  と定義する。これらの関係は図 5.1 に、模式的に示されている。なお、特に指定がなければ本章で述べる角度差とは、真の SN 方向と再構成された方向との角度差  $\Delta\theta_{SN}$  のことを示す。

$$\cos \Delta\theta_{SN} = \hat{d}_{SN}^{\text{true}} \cdot \hat{d}_{SN}^{\text{recon}} \quad (5.1)$$

SN ニュートリノバーストにおいて得られる角度差  $\Delta\theta_{SN}$  は、ES 事象と非 ES 事象の数やそれらの運動方向分布に依存するため、個々のバーストごとに異なる値を示す。そのため、本

研究分野では、多数の MC データに対して同一の解析手法を適用し、各 MC 毎に得られる角度差  $\Delta\theta_{SN}$  の分布を用いて、方向再構成精度を統計的に評価する手法が一般的に用いられている。図 5.2 には、SN までの距離 10kpc における 1000 個の MC データセットに対して DBSCAN-fitter を適用して得られた  $\Delta\theta_{SN}$  分布を示す。

本研究では、この  $\Delta\theta_{SN}$  分布において  $\Delta\theta_{SN} = 0^\circ$  からの累積分布が全体の 68% に達する角度を角度分解能 ( $1\sigma$ ) と定義し、他の方向再構成手法との比較に用いた。

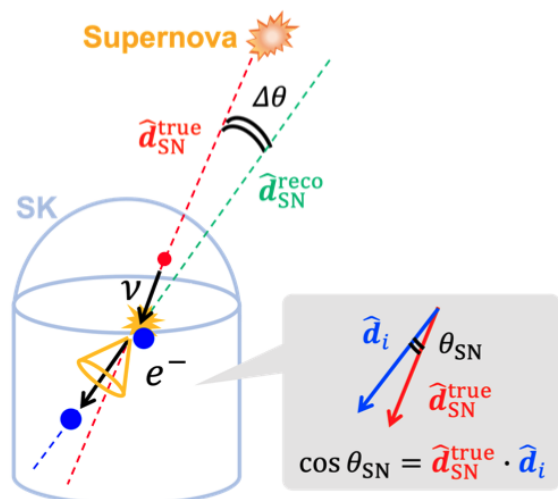


図 5.1: 真の SN 方向と再構成方向間の  $\Delta\theta_{SN}$

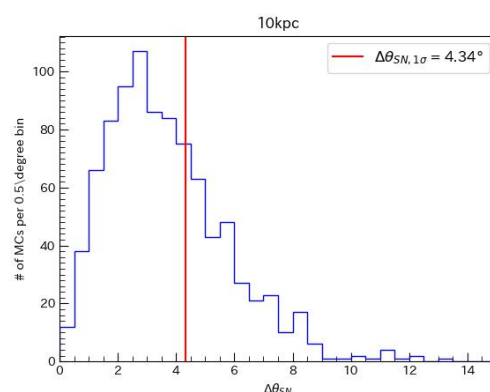


図 5.2: SN 距離 10kpc における 919 個の MC に対し DBSCAN-fitter を適用して得た  $\Delta\theta_{SN}$  分布

## 5.2.2 Failure rate

DBSCAN-fitter は、事象数や ES 事象および非 ES 事象の運動方向分布に依存して、クラスタ抽出に失敗する場合がある。また、クラスタ抽出に成功した場合でも、再構成された方向が真の SN 方向と大きく乖離することがある。そこで本研究では、方向再構成手法の安定性を評価するため、二種類の Failure-rate を導入した。一つは DBSCAN によるクラスタ抽出の成否に関する指標、もう一つは先行研究において一般的に用いられる方向再構成の失敗率である。

まず、クラスタ抽出の成否に関する Failure-rate を定義する。本研究において DBSCAN に求める理想的な分類は、ES 事象が集中したクラスタ領域と、それ以外の事象からなる領域の二つに分類されることである。この要件に基づき、以下の二つを失敗と定義した。

1. DBSCAN が SN 事象を二種類の領域に分類できない場合
2. DBSCAN により抽出されたクラスタ領域を用いて再構成した方向と、真の SN 方向との角度差  $\Delta\theta_{SN}$  が  $90^\circ$  以上となる場合

本論文では、他で定義されている Failure-rate との混同を避けるため、この量を Not-finding-rate と呼ぶ。ここで、これらの条件のいずれかを満たす MC データの数を  $N_{fail}$ 、方向再構成に成功した MC データの数を  $N_{good}$  とすると、Not-finding-rate は式 (5.2) で定義される。

$$R_{fail} = \frac{N_{fail}}{N_{fail} + N_{good}} \quad (5.2)$$

次に、先行研究で一般的に用いられる Failure-rate の定義について述べる。これらの研究では、Failure-rate は式 5.1 で定義される  $\cos \Delta\theta_{SN}$  分布を用いて Failure-rate を推定する。図 5.3 に、SN 距離 20kpc における約 1000 個の MC データ群に対して DBSCAN-fitter を適用した際に得られた  $\cos \Delta\theta_{SN}$  分布を示す。この分布から、 $\cos \Delta\theta_{SN} = 1$  付近に顕著なピークを持ち、それ以外の領域では一様に分布する様子が確認できる。この分布形状に基づき、 $\cos \Delta\theta_{SN} \in [-1, 0]$  の範囲に含まれる MC データを方向再構成に失敗したデータであると仮定する。この仮定のもとで、当該領域における 1 ビンあたりの平均失敗データ数  $\bar{N}_{fail,bin}$  を求め、全ビン数  $n_{bins,total}$  を用いて、再構成に失敗した総データ数  $N_{fail}$  を

$$N_{fail} = \bar{N}_{fail,bin} \times n_{bins,total} \quad (5.3)$$

と推定する。全データ数から  $N_{fail}$  を差し引くことで、方向再構成に成功した数  $N_{good}$  を求め、式 5.2 を用いて Failure-rate を算出する。本研究では、この Failure-rate と、前述の Not-finding-rate を併用することで、方向再構成手法の安定性を多角的に評価する。

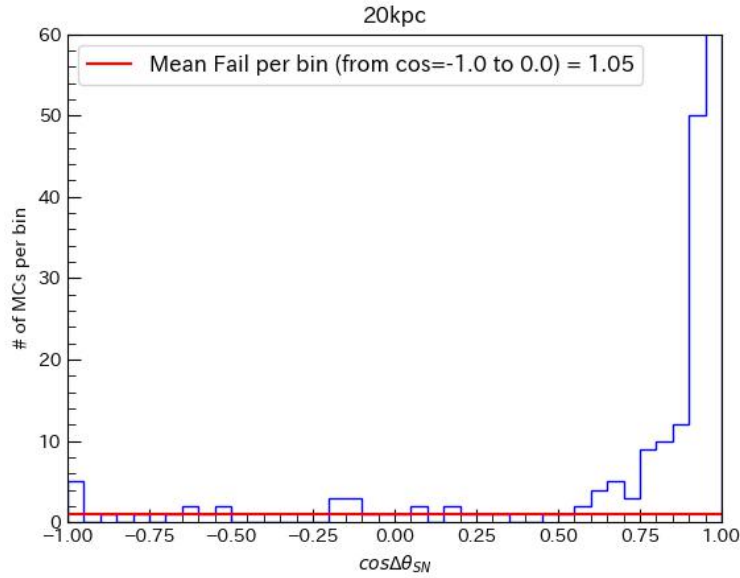


図 5.3: 20kpc の MC データ群に対して DBSCAN-fitter を適用して得た  $\cos \Delta\theta_{SN}$  分布。 $\cos \Delta\theta_{SN}=1$  付近にピークを持ち、それ以外の領域では一様分布が見られる。

### 5.2.3 Reconstruction time

本研究では、DBSCAN-fitter の処理時間を、SN ニュートリノバースト発生時におけるリアルタイム解析への適用可能性を評価する目的で測定した。本評価では、4.2 節で述べた解析フローに基づき、再構成済みの事象に対して各種事象情報を用いた選別を行ってから、最終的に一つの再構成 SN 方向を抽出するまでの処理を対象とした。したがって、IBD タグ付けによる中性子事象の除去や、WIT による SN ニュートリノ事象の再構成処理は、本処理時間の評価に含まれていない。本解析は Apple M1(8-core CPU)、8GB メモリを搭載した macOS 環境上において、Anaconda 環境の Jupyter Notebook を用いて行われた。

## 5.3 DBSCAN パラメータ推定

本研究における SN 方向再構成における最初の重要な過程は、ES 事象が集中したクラスタ領域を適切に抽出することである。そのためには DBSCAN パラメータを各 SN 事象に対して適切に推定しなければならない。また、真の SN ニュートリノバーストを観測時に、真の SN 距離は知ることができないため、パラメータ推定は観測事象数に基づいて行われる必要がある。本節では、`eps` および `min_samples` の推定式を構築した過程とその結果について述べる。なお、パラメータ推定は各 SN 距離ごとに 300 個の訓練データを用いて行った。なお、以下では 6MeV より高いエネルギー事象を用いた場合を上付き文字 `high`、全エネルギー事象を用いた場合を上付き文字 `all` として区別して記述する。

DBSCAN 解析において、これら二つのパラメータのうち `eps` を不適切な値に設定した場合、全ての事象がノイズとして分類される、あるいは全事象が単一クラスタとして誤って抽出されるなど、クラスタ抽出そのものが破綻する可能性がある。このため、本研究では `min_samples` に先立って `eps` の最適化を優先的に行った。本研究では、4.2.2 項で述べた通り、SN 事象から直接得られる事象数や各事象の局所密度、運動方向分布の構造と、方向再構成に最も適した `eps` との間に十分強い相関を見出すことができなかったため、`eps` および `min_samples` の様々な組み合わせに対して DBSCAN 解析を行い、それぞれから得られる方向再構成結果の角度差  $\Delta\theta_{SN}$  を評価することで、DBSCAN パラメータ推定式を構築する方針とした。

### 5.3.1 `eps` の最適化 (`Up_eps`、`Middle_eps`、`Bottom_eps`)

本研究では 3 種類の `eps` を用いて解析を行うが、まず中間値に相当する `Middle_eps` の導出過程について述べる。`Middle_eps` には、多数の SN ニュートリノバースト事象に対して、ES 事象が集中したクラスタ領域を安定して抽出でき、かつ高い精度で良好な方向再構成精度を与える代表的な `eps` を選択する方針とした。以下に、その具体的な抽出手順を示す。本手順は煩雑であるため、ここでは代表例として SN 距離 10kpc の場合に限定して説明する。なお、パラ

メータ推定においては、DBSCAN により抽出されたクラスタに含まれる各事象の運動方向を単純平均することで、再構成 SN 方向を決定している。

1. 各 SN 距離毎に角度差  $\Delta\theta_{SN}$  の閾値を設定する。本閾値は「良好な方向再構成が行われやすいか」を判定するための作業基準として設定した。閾値の設定にあたっては、本研究分野で良好な結果を示す SNWATCH の現行再構成手法による角度分解能 [17] を参照し、 $0.5^\circ$  刻みに粗視化した値を用いた。なお、Middle\_eps の導出では、各 eps に対して良好な再構成結果を示した SN ニュートリノバースト事象数に着目している。各 eps と良好な結果を示す SN 事象数の関係は、閾値を参照値から  $\pm 0.25$  程度変動させてもほとんど変化しないことを確認した。以上より、本研究では上記の閾値を採用した。
2. 1 つの訓練データに対し、多数の eps と min\_samples の組み合わせについて DBSCAN 解析を行う。図 5.4 に、10kpc の訓練データに対して eps = 0.4 に固定し、min\_samples を変化させながら DBSCAN を適用した際に得られる角度差  $\Delta\theta_{SN}$  を示す。この分布から、設定した角度差  $\Delta\theta_{SN}$  の閾値よりも良好な角度差を示す min\_samples の連続した領域が得られる。
3. 閾値よりも良好な分解能を示す min\_samples の連続領域の幅が 10 以上である場合、その訓練データは当該 eps に対して「良好な角度差領域を持つデータ」と定義する。これは、min\_samples の選択に対して結果が過度に敏感でない、すなわちパラメータに対して安定な挙動を示すことを要請するためである。
4. 同一の eps に対して他の訓練データにも同様の操作を行い、「良好な角度差領域を持つデータ」の数をカウントする。
5. 以上の操作を eps を変えて繰り返すことで、各 eps に対する「良好な角度差領域を持つデータ」の数を得る。その結果を図 5.5 に示す。この分布から、良好な角度差を示す条件が特定の eps の値に集中していることが分かる。
6. 最後に、「良好な角度差領域を持つデータ」数が多い上位 10 個の eps の平均値を算出し、その値を SN 距離 10kpc における eps の代表値として用いる。ここで平均 eps を算出するのは、統計的な揺らぎや偶発的なピークの影響を低減するためである。

表 5.1: Middle\_eps のパラメータ最適化において使用した角度差  $\Delta\theta_{SN}$  の閾値

超新星までの距離	角度差 $\Delta\theta_{SN}$ の閾値
4kpc	1.50
6kpc	2.00
8kpc	3.00
10kpc	3.50
12kpc	4.50
14kpc	5.50
16kpc	6.50
18kpc	7.50
20kpc	9.00

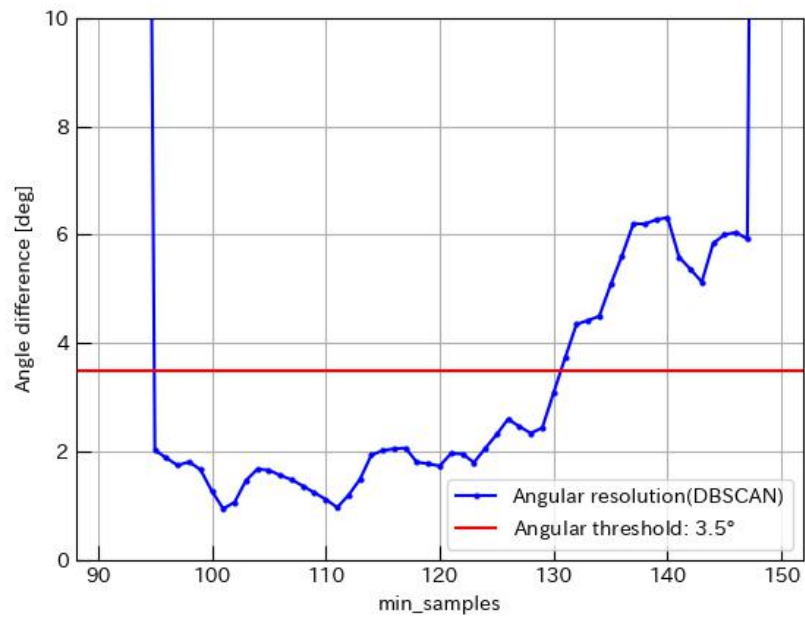


図 5.4: 10kpc の訓練データに対し、 $\text{eps} = 0.4$  に固定して DBSCAN を適用した際の角度差  $\Delta\theta_{SN}$  ( $\text{min\_samples}$  依存性)。良好な角度差を示す  $\text{min\_samples}$  の連続領域が見られる。

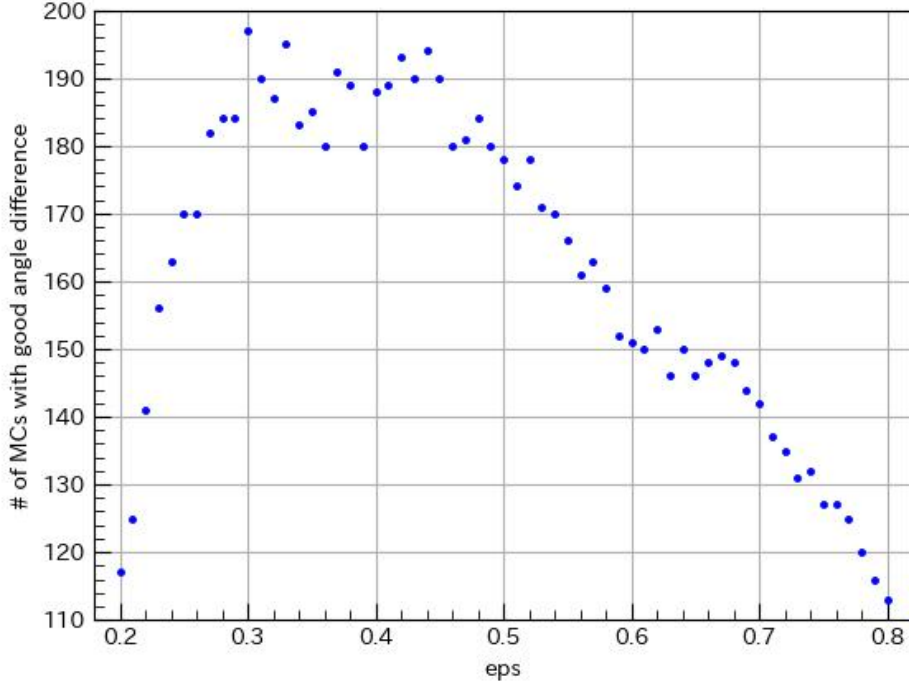


図 5.5: 10kpc の訓練データ 300 個の用いて得られた、各 eps に対する「良好な角度差領域をもつデータ」数。多くの訓練データに対して良好な角度差  $\Delta\theta_{SN}$  を得やすい eps 領域がある。

以上の手順に従って各 SN 距離毎に抽出される平均 eps と平均事象数との関係から、Middle\_eps の定義式を構築する。

高エネルギー事象に対して、この手順によって抽出された平均 eps と平均事象数との関係を図 5.6 に示す。図 5.6 は両対数グラフであり、この分布から、事象数と平均 eps の間にべき乗的な関係が成り立つことが分かる。そこで、Middle\_eps の推定には、以下のべき乗関数を用いてフィッティングを行った。

$$y = A \times x^k \quad (5.4)$$

事象数を  $N_{events}$  とすると、フィッティングから得られた高エネルギー事象に対する Middle\_eps の推定式は、次式で表される。

$$\text{Middle\_eps}^{\text{high}} = 2.188 \times N_{events}^{-0.234} \quad (5.5)$$

次に、Middle\_eps を基準として設定する補助的なパラメータ (Up\_eps、Bottom\_eps) の推定式構築手順について述べる。なお、これまでと同様に SN 距離 10kpc の場合を例として説明する。

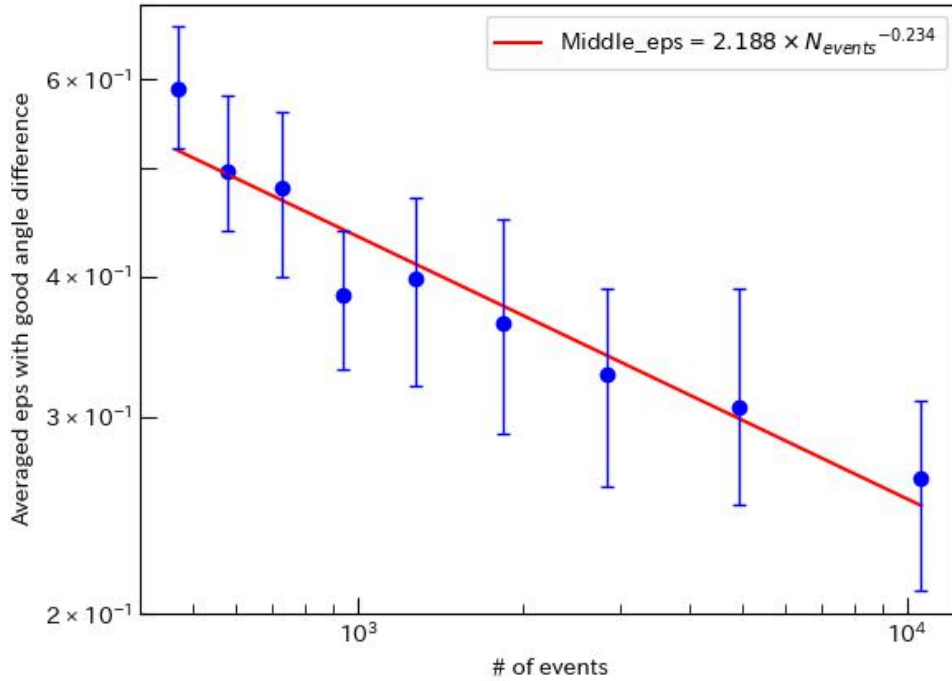


図 5.6: 高エネルギー事象に対して、5.2.1 項の手順に従って抽出される平均 eps と各距離に対応する平均事象数の関係 (この分布における縦線は、「良好な角度差領域を持つデータ数」が多い 10 個の eps の範囲を示す)

1. 各訓練データごとに Middle\_eps を算出する。
2. Up\_eps の導出では、Middle\_eps + 0.05 を下限とし、そこから 0.2 を加えた値を上限として eps の探索範囲を設定する。一方、Bottom\_eps の導出では、Middle\_eps - 0.05 を上限とし、そこから 0.2 を引いた値を下限として探索範囲を設定する。なお、ここで用いた eps の刻み幅 0.05 は Middle\_eps から 0.1 程度離れると良好な角度差を持つ訓練データ数が急激に減少するという観測事実に基づき設定した。また、探索幅 0.2 は Bottom\_eps 構築のために用いる eps 探索範囲が負になることへの回避と、過大クラスタや有効例の少ない高 eps 領域の除外を目的として設定した。
3. 設定した eps の探索範囲内で、多数の min\_samples と組み合わせて DBSCAN を適用し、各訓練データに対して最も角度差  $\Delta\theta_{SN}$  が良好となる eps を抽出する。
4. 過程 3 で抽出した eps について、各訓練データ間で平均値を算出する。

高エネルギー事象に対して、以上の手順に従って抽出した各 SN 距離における Middle\_eps より大きい、または小さい平均 eps と、平均事象数との関係を図 5.7 に示す。この分布に対し

ても、式 (5.4) を用いてフィッティングを行い、Up\_eps と Bottom\_eps の推定式を構築した。

$$\text{Up\_eps}^{\text{high}} = 1.819 \times N_{\text{events}}^{-0.168} \quad (5.6)$$

$$\text{Bottom\_eps}^{\text{high}} = 2.109 \times N_{\text{events}}^{-0.283} \quad (5.7)$$

三種類の eps の推定式構築手順を、全エネルギー事象に対しても適用した。抽出された各平均 eps と平均事象数との関係を図 5.8 に示す。この分布に対しても、同様に式 (5.4) を用いてフィッティングを行い、以下の推定式を得た。

$$\text{Up\_eps}^{\text{all}} = 2.151 \times N_{\text{events}}^{-0.187} \quad (5.8)$$

$$\text{Middle\_eps}^{\text{all}} = 2.863 \times N_{\text{events}}^{-0.265} \quad (5.9)$$

$$\text{Bottom\_eps}^{\text{all}} = 2.914 \times N_{\text{events}}^{-0.320} \quad (5.10)$$

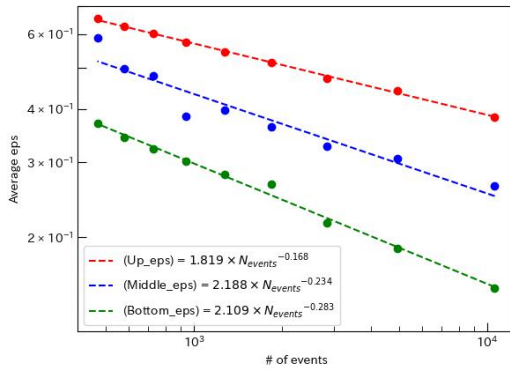


図 5.7: 高エネルギー事象に対して、上記の手順に従って抽出される Middle\_eps より高い、または低い平均 eps と各距離に対応する平均事象数の関係

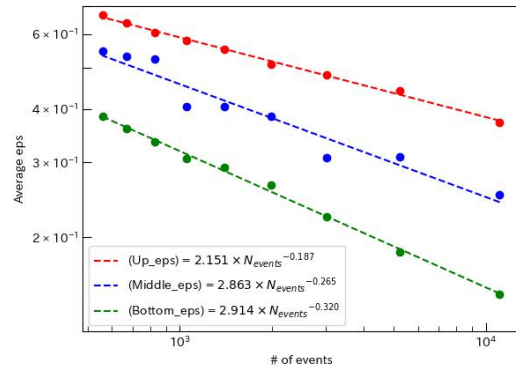


図 5.8: 全エネルギー事象に対して、上記の手順に従って抽出される Middle\_eps より高い、または低い平均 eps と各距離に対応する平均事象数の関係

### 5.3.2 各 eps に対する min\_samples の推定式構築

各 eps に対応する min\_samples の推定式の構築について述べる。DBSCAN によるクラスタ抽出において、eps と min\_samples の組み合わせが結果に大きな影響を与えるため、min\_samples は、事象数だけでなく eps との関係も同時に考慮した形で構築する必要がある。

まず、eps を大きく設定すると、クラスタ探索に用いる近傍領域が広がるため、小さな min\_samples を与えた場合には、距離の離れたデータ点同士が同一クラスタとして誤って統合される可能性が高くなる。従って、適切なクラスタ抽出を行うために、eps の増大に応じ

て `min_samples` も増加させる必要がある。一方、事象数が増加すると、単位空間あたりのデータ点密度が高くなるため、偶然形成される高密度領域の抽出を防ぐには、より大きな `min_samples` を設定する必要がある。このことから、`min_samples` は `eps` および事象数の双方と正の相関を持つと考えられる。

以上の考察より、`min_samples` は `eps` および事象数の増大に伴って単調に増加する量であると考えられる。そこで本研究では、近傍体積に関する `eps` と、事象数  $N_{\text{events}}$  の積を一つの指標とし、`min_samples` が  $(\text{eps} \times N_{\text{events}})$  に対してべき乗則的に依存するという仮定のもとで、`min_samples` の推定式を構築した。

次に、各 `eps` に対する `min_samples` の推定式を構築する具体的な手順について述べる。図 5.9 に `eps` を 0.4 に固定した場合における、100 個の訓練データと、Middle\_eps 抽出過程 2 から得られた「良好な角度差を持つ `min_samples` の範囲」との関係を示す。この分布から、同一 `eps` に対して良好な角度差  $\Delta\theta_{SN}$  を与える `min_samples` の値は、同一距離においては訓練データ間でおおよそ一致していることが分かる。この性質を利用し、以下の手順に従って `min_samples` を推定式を構築した。なお、以下では代表例として SN 距離 10kpc、Middle\_eps に対する `min_samples` の推定式構築例を示す。

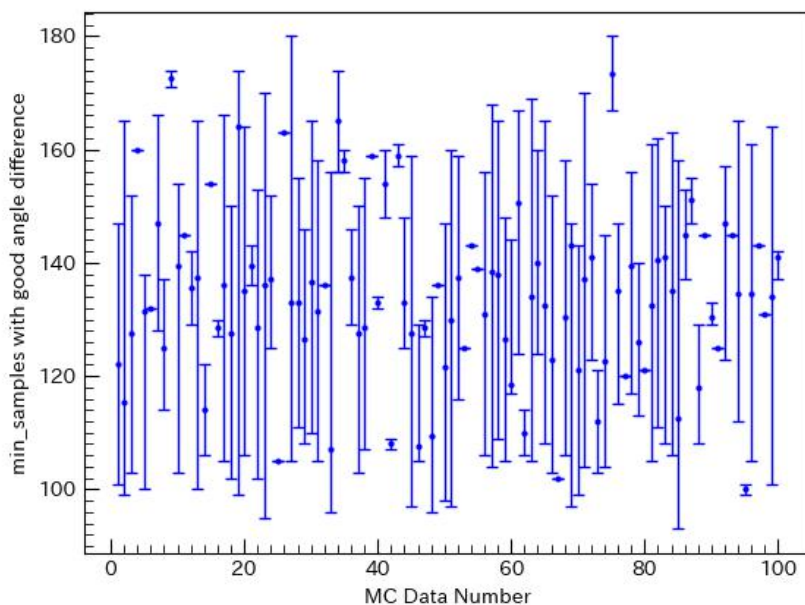


図 5.9: 10kpc、100 個の訓練データに対し `eps` を 0.4 に固定した場合における、各訓練データのラベルと、middle\_eps 抽出過程 2 から得られた「良好な角度差を持つ `min_samples` の範囲」との関係。良好な角度差を示す `min_samples` の値は同一距離でおおよそ一致する。

1. 各訓練データごと Middle\_eps を算出し、その Middle\_eps に対して良好な角度差  $\Delta\theta_{SN}$  を示す min\_samples の範囲を求める。
2. 各訓練データについて、過程 1 で得られた min\_samples の範囲から中央値を求める。
3. 過程 2 で得られた各訓練データの min\_samples の中央値を用いて平均値を算出する。

最終的に、以上の手順に従って各 SN 距離ごとに求めた平均 min\_samples、対応する Middle\_eps、および平均事象数を用いて分布を作成し、これらの関係に対してフィッティングを行うことで min\_samples の推定式を構築した。

Bottom\_eps と Up\_eps に対応する min\_samples の推定式についても同様の手順に従う。高エネルギー事象に対して、以上の手順に従って求めた各 SN 距離ごとに算出した平均 min\_samples と、各 eps および平均事象数との関係を表す分布を図 5.10 に、全エネルギー事象に対して作成したこれらの関係を示す分布を 5.11 に示す。これらの分布は両対数グラフで表されており、本項で仮定した通り、 $(eps \times N_{events})$  と平均 min\_samples の間にべき乗則的な関係が成り立っていることが確認できる。そこで、これまでと同様に、式 5.4 に表されるべき乗関数を用いてフィッティングを行い、三種類の eps に対応する min\_samples の推定式を構築した。得られた各 eps に対応する min\_samples の推定式を表 5.2 にまとめた。

表 5.2: 各 eps に対応する min\_samples の推定式

eps	高エネルギーにおける min_samples	全エネルギーにおける min_samples
Up_eps	$0.532 \times (eps \times N_{events})^{0.851}$	$0.541 \times (eps \times N_{events})^{0.850}$
Middle_eps	$0.693 \times (eps \times N_{events})^{0.779}$	$0.778 \times (eps \times N_{events})^{0.762}$
Bottom_eps	$1.181 \times (eps \times N_{events})^{0.641}$	$1.241 \times (eps \times N_{events})^{0.638}$

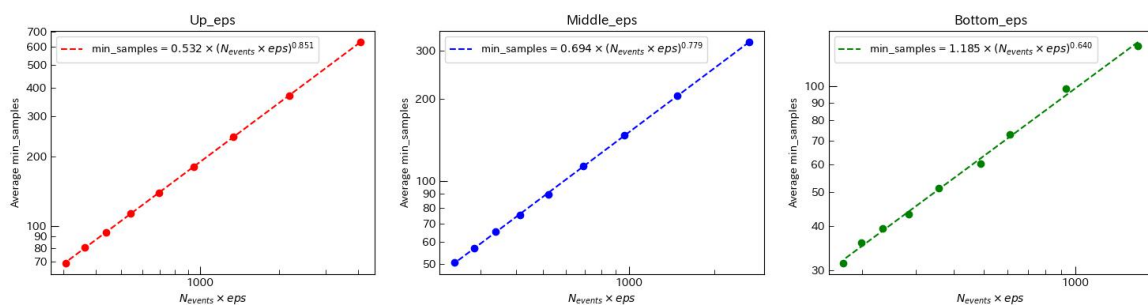


図 5.10: 高エネルギー事象に対して、5.2.2 項の手順に従って抽出される平均 min\_samples と各距離に対応する平均事象数の関係 (左:Bottom\_eps に対応、中央:Middle\_eps に対応、右:Up\_eps に対応)。それぞれべき乗則的な関係が見られる。

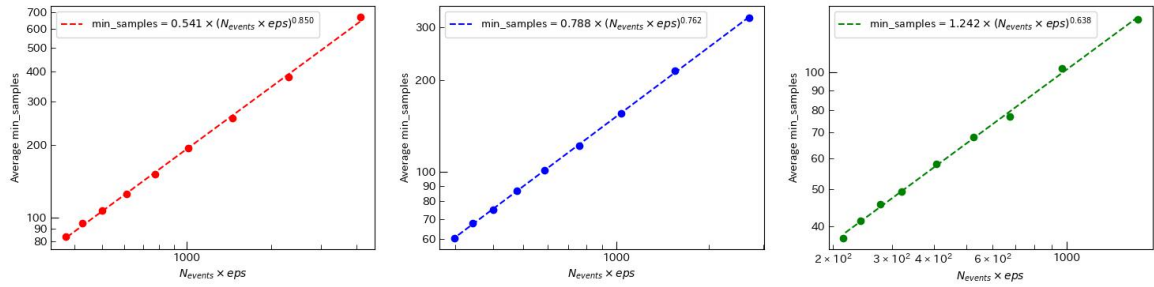


図 5.11: 全エネルギー事象に対して、5.2.2 項の手順に従って抽出される平均 min\_samples と各距離に対応する平均事象数の関係 (左:Bottom\_eps に対応、中央:Middle\_eps に対応、右:Up\_eps に対応)。それぞれべき乗則的な関係が見られる。

以上の手順により 3 種類の eps (Bottom\_eps、middle\_eps、Up\_eps) とそれぞれに対応する min\_samples の推定式が構築される。本研究では、各 eps に対して得られた min\_samples を中心とし、その値を 1.075 倍および 0.925 倍した二種類の min\_samples を追加で用いることで探索範囲を拡張した。これらの定数の妥当性および最終的な性能への影響については、第 6 章で述べる。

### 5.3.3 DBSCAN パラメータ表

5.2.1 項および 5.2.2 項で構築した、高エネルギー事象と全事象エネルギー事象に対する 9 組の DBSCAN パラメータセット推定式の一覧を表 5.3 と図 5.4 にまとめた。これらのパラメータセットは各 SN 距離における観測事象数に基づいて決定され、後続の解析において一貫して用いられる。

表 5.3: 高エネルギー事象に対して用いる 9 組のパラメータセット

eps	eps の推定式	min_samples の推定式	追加 min_samples
Up_eps	$1.819(N_{\text{events}})^{-0.168}$	$0.532(Up\_eps \times N_{\text{events}})^{0.851}$	$\times 0.925, 1.075$
Middle_eps	$2.188(N_{\text{events}})^{-0.234}$	$0.693(Middle\_eps \times N_{\text{events}})^{0.779}$	$\times 0.925, 1.075$
Bottom_eps	$2.109(N_{\text{events}})^{-0.283}$	$1.181(Bottom\_eps \times N_{\text{events}})^{0.641}$	$\times 0.925, 1.075$

表 5.4: 全エネルギー事象に対して用いる 9 組のパラメータセット

eps	eps の推定式	min_samples の推定式	追加 min_samples
Up_eps	$2.151(N_{\text{events}})^{-0.187}$	$0.5341(Up\_eps \times N_{\text{events}})^{0.850}$	$\times 0.925, 1.075$
Middle_eps	$2.863(N_{\text{events}})^{-0.265}$	$0.778(Middle\_eps \times N_{\text{events}})^{0.762}$	$\times 0.925, 1.075$
Bottom_eps	$2.914(N_{\text{events}})^{-0.320}$	$1.241(Bottom\_eps \times N_{\text{events}})^{0.638}$	$\times 0.925, 1.075$

## 5.4 k-Nearest Neighbors 再構成法のパラメータ推定

kNN 方向再構成手法において設定すべきパラメータは、Fraction の計算に必要な  $\cos$  閾値、kNN 本来のパラメータ  $n\_neighbors$ 、および式 4.1 の重み付け関数に含まれるパラメータ  $a, b$  の計 4 種類である。本節では、観測された事象数に基づいて、これらのパラメータ推定式の構築について述べる。

まず、Fraction の計算に用いる  $\cos$  閾値について述べる。この閾値は一意には決定せず、複数の値を用いた場合の角度分解能を比較することで、適切な値を判断する方針とした。これは、DBSCAN-fitter による最終的な角度分解能が、kNN 方向再構成法のみならず、DBSCAN パラメータや結果選別に用いる Estimator など、複数の要因の組み合わせによって決定されるためである。本研究では、 $\cos$  閾値候補として 0.85、0.90、0.95 の 3 通りを採用した。これら閾値による角度分解能への影響については、第 6 章で評価する。

次に、 $n\_neighbors$  の最適化について述べる。本研究における kNN は、各事象周辺の局所密度を評価する役割に加え、DBSCAN によって抽出されたクラスタ内に混在するバックグラウンド事象を抑制する役割を担う。そのため、ES 事象と非 ES 事象が分離される  $n\_neighbors$  の値を採用する方針とした。

図 5.12 に、DBSCAN によって抽出されたクラスタに対して kNN を適用した際に得られる MeanKdistance 分布を示す。この分布から、 $n\_neighbors$  の増加に伴い、ES 事象と非 ES 事象がそれぞれ集中する領域が分離されていく様子が確認できる。この分離を定量的に評価するため、MeanKdistance に閾値を設定し、その閾値以下に含まれる事象数を用いて ES 事象の Significance(式 3.1) を算出した。その後、各  $n\_neighbors$  毎に Significance が最大となる値を抽出することで、その最大値と  $n\_neighbors$  の関係を調べた。その結果を図 5.13 に示す。この分布から、 $n\_neighbors$  の増加に伴って最大 Significance は増大し、最終的に飽和する傾向が確認できる。

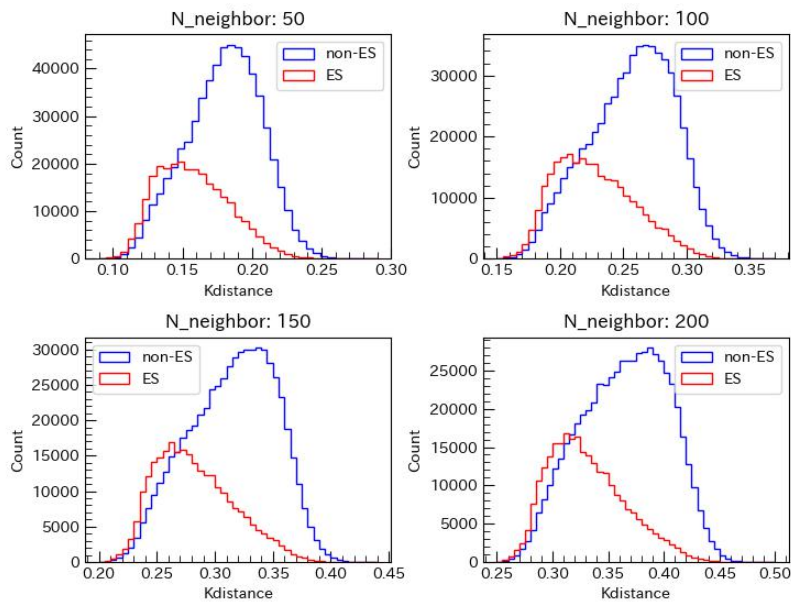


図 5.12: 10kpc、300 個の訓練データにおいて、DBSCAN 分類後のクラスタに対し、k-Nearest Neighbors を適用して得られる MeanKdistance 分布。n\_neighbors が大きくなると ES 事象と非 ES 事象の集中領域が分離する。

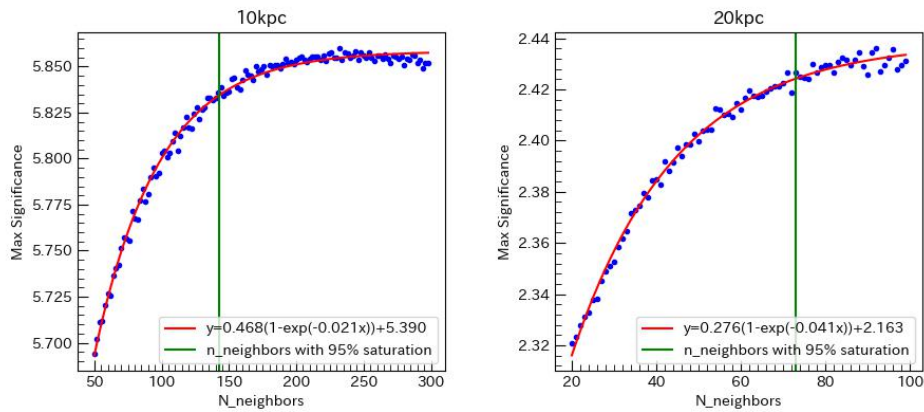


図 5.13: 300 個の訓練データに対して得られた n\_neighbors と最大 Significance の分布 (左:SN 距離 10kpc、右:SN 距離 20kpc)。n\_neighbors の増加に伴って、Significance は単調に増加していき、最終的に飽和する。

本研究では、処理時間とのトレードオフを考慮し、飽和 Significance の 95%に相当する値を示す n\_neighbors を各 SN 距離毎に代表値として抽出した。高エネルギー事象および全エネルギー事象に対して、上記の手順に従って抽出した各 SN 距離毎の n\_neighbors と平均事象数との関係を表す分布を、図 5.14 および図 5.15 に示す。いずれの分布も両対数グラフで表されて

おり、DBSCAN パラメータ推定の場合と同様に、事象数と  $n\_neighbors$  の間にべき乗則的な関係が成り立つことが分かる。そこで、式 5.4 のべき乗関数を用いてこれらの分布に対してフィッティングを行い、 $n\_neighbors$  の推定式を構築した。高エネルギー事象および全エネルギー事象に対する推定式は、それぞれ次式で表される。

$$n\_neighbors^{high} = 0.690 \times N_{events}^{0.689} \quad (5.11)$$

$$n\_neighbors^{all} = 1.378 \times N_{events}^{0.619} \quad (5.12)$$

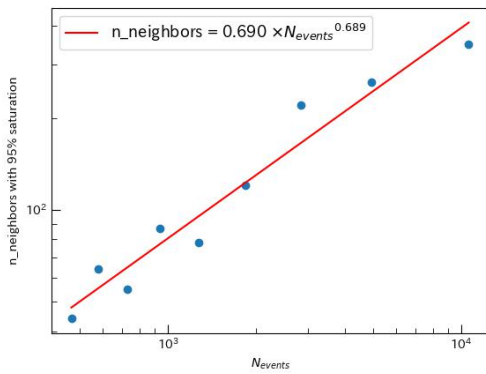


図 5.14: 高エネルギー事象において飽和 Significance の 95%を示す  $n\_neighbors$  と平均事象数の関係

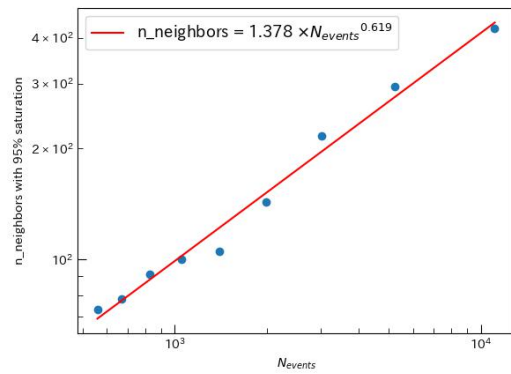


図 5.15: 全エネルギー事象において飽和 Significance の 95%を示す  $n\_neighbors$  と平均事象数の関係

最後に、重み付け関数  $F_i(\text{MeanKdistance})$ (式 4.1) に含まれるパラメータ  $(a, b)$  の推定式構築について述べる。式 5.13 で定義される Fraction を用いて、MeanKdistance と Fraction の関係分布を作成する。作成した分布を図 5.16 に示す。

$$\text{Fraction} = \frac{\# \text{ of events with cosine} \geq \text{threshold in the bin}}{\text{total} \# \text{ of events in the bin}} \quad (5.13)$$

作成した図 5.16 の分布に対して、式 4.1 で表される関数を用いてフィッティングを行い、各 SN 距離毎の  $(a, b)$  を抽出した。

高エネルギー事象および全エネルギー事象に対して、各 SN 距離と各 cos 閾値毎に抽出した重み付け関数パラメータ  $a, b$  と平均事象数との関係を図 5.17、図 5.18、図 5.19、図 5.18 に示す。これらの分布から、事象数と各パラメータ  $(a, b)$  の間にもべき乗則的な関係が成り立つことが確認できる。そこで、式 5.4 のべき乗関数を用いてフィッティングを行い、各パラメータの推定式を構築した。高エネルギー事象および全エネルギー事象に対する重み付け関数パラメータ  $a, b$  の推定式をそれぞれ表 5.5 と表 5.6 にまとめた。

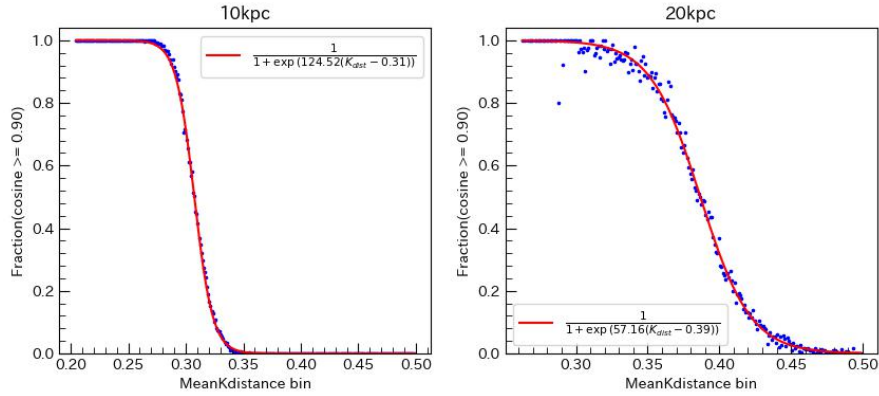


図 5.16: 10kpc において、MeanKdistance と Cosine 値の閾値を 0.9 に設定したときの Fraction 分布。青点が各ビンの Fraction を、赤線がそれらをフィッティングして得られる直線を示す。

表 5.5: 高エネルギー事象に対する重み付け関数パラメータの推定式

cos 閾値	n_neighbors	パラメータ $a$	パラメータ $b$
0.85	$0.690(N_{\text{events}})^{0.689}$	$2.493(N_{\text{events}})^{0.501}$	$0.809(N_{\text{events}})^{-0.132}$
0.90	$0.690(N_{\text{events}})^{0.689}$	$3.510(N_{\text{events}})^{0.466}$	$0.790(N_{\text{events}})^{-0.136}$
0.95	$0.690(N_{\text{events}})^{0.689}$	$2.187(N_{\text{events}})^{0.526}$	$0.786(N_{\text{events}})^{-0.147}$

表 5.6: 全エネルギー事象に対する重み付け関数パラメータの推定式

cos 閾値	n_neighbors	パラメータ $a$	パラメータ $b$
0.85	$1.378(N_{\text{events}})^{0.619}$	$2.481(N_{\text{events}})^{0.500}$	$1.168(N_{\text{events}})^{-0.170}$
0.90	$1.378(N_{\text{events}})^{0.619}$	$3.207(N_{\text{events}})^{0.474}$	$1.178(N_{\text{events}})^{-0.177}$
0.95	$1.378(N_{\text{events}})^{0.619}$	$1.652(N_{\text{events}})^{0.556}$	$1.232(N_{\text{events}})^{-0.193}$

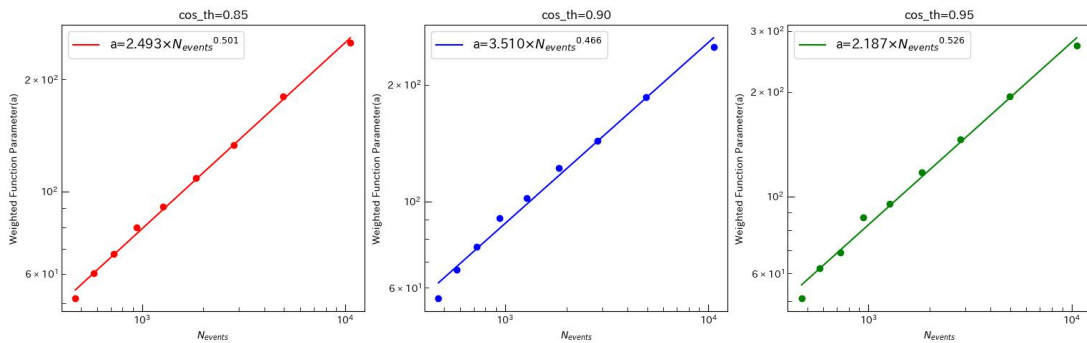


図 5.17: 高エネルギー事象における、抽出した重み付け関数パラメータ  $a$  と平均事象数の関係 (左:cos 閾値=0.85、中央:cos 閾値=0.90、右:cos 閾値=0.95)

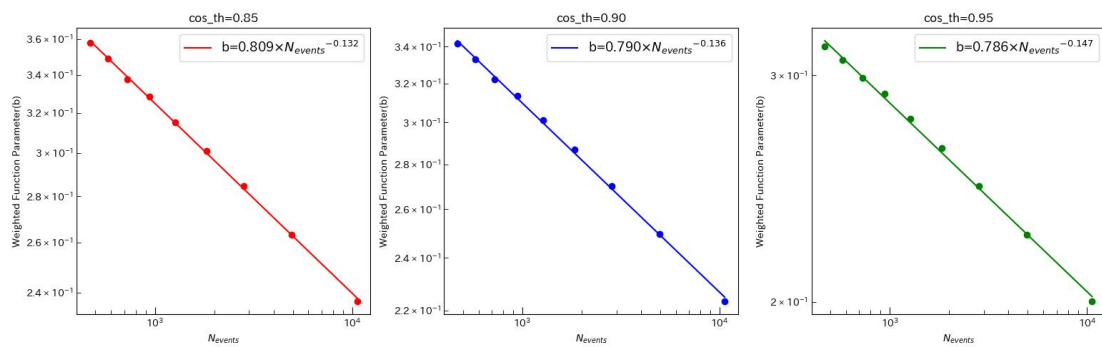


図 5.18: 高エネルギー事象における、抽出した重み付け関数パラメータ  $b$  と平均事象数の関係 (左: $\cos$  閾値=0.85、中央: $\cos$  閾値=0.90、右: $\cos$  閾値=0.95)

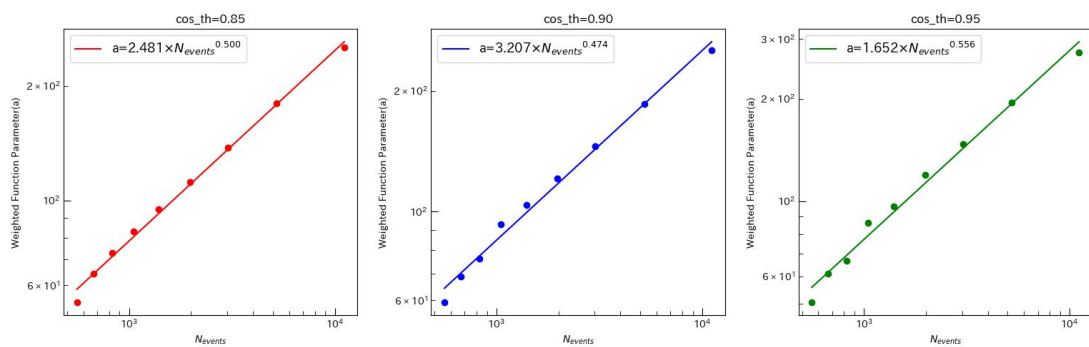


図 5.19: 全エネルギー事象における、抽出した重み付け関数パラメータ  $a$  と平均事象数の関係 (左: $\cos$  閾値=0.85、中央: $\cos$  閾値=0.90、右: $\cos$  閾値=0.95)

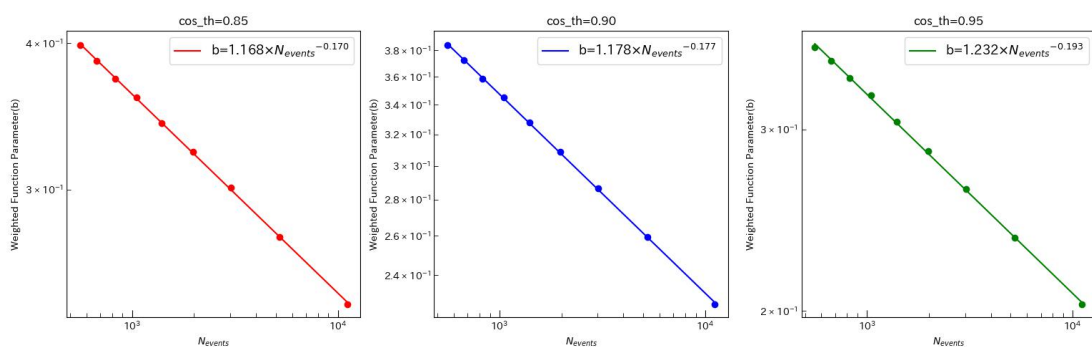


図 5.20: 全エネルギー事象における、抽出した重み付け関数パラメータ  $b$  と平均事象数の関係 (左: $\cos$  閾値=0.85、中央: $\cos$  閾値=0.90、右: $\cos$  閾値=0.95)

## 5.5 Minimum Distance Estimator のパラメータ推定

本節では、Minimum Distance Estimator(MDE) において距離和の計算に用いる事象数  $N_{\text{MDE}}$  の推定方法について述べる。MDE による評価は使用する事象数  $N_{\text{MDE}}$  に基づくため、この値がわずかに変化するだけでも、角度分解能が大きく変化する可能性がある。そのため本研究では、分解能が安定して高く保たれる  $N_{\text{MDE}}$  の連続領域を抽出する方針を採用した。

まず、各 SN 距離毎の訓練データに対して、複数の  $N_{\text{MDE}}$  を設定して MDE を適用し、角度分解能を算出した。SN 距離 10kpc の訓練データ 300 個に対する  $N_{\text{MDE}}$  と角度分解能の関係を図 5.21 に示す。この分布から、角度分解能は  $N_{\text{MDE}}$  に対して非単調かつ不連続に変化しており、わずかな  $N_{\text{MDE}}$  の変化によっても再構成性能が大きく変化的ことが確認できる。次に、局所的な揺らぎを抑制する目的で隣接する 5 点の移動平均を適用した結果を図 5.22 に示す。しかし、平均化後も分解能の揺らぎが残っており、単一の最小値に基づくパラメータ選択は不安定であることが分かる。

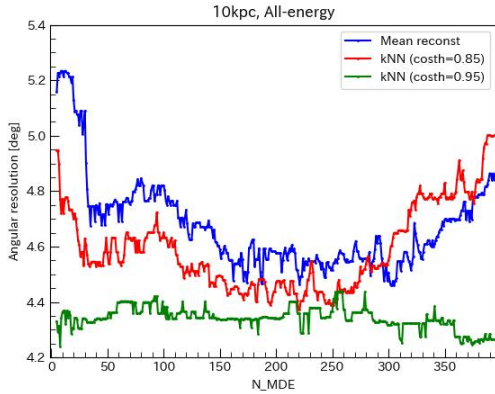


図 5.21: SN 距離 10kpc の訓練データ 300 個に対する、全エネルギー事象での  $N_{\text{MDE}}$  と角度分解能の分布

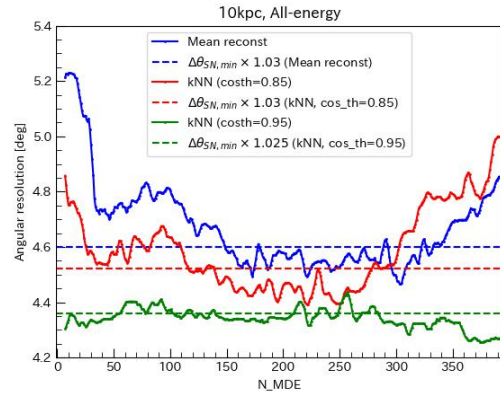


図 5.22: SN 距離 10kpc の訓練データ 300 個に対する、全エネルギー事象での  $N_{\text{MDE}}$  と角度分解能の分布に対し隣接する 5 点の移動平均を適用した結果

そこで本研究では、分解能の最小値に対して一定割合以内の値を示す  $N_{\text{MDE}}$  の連続領域を有効な範囲として定義した。具体的には、分解能の最小値にその値の 3%を加えた値を閾値とし、その値を下回る  $N_{\text{MDE}}$  の領域を抽出した。抽出された領域が複数存在する場合には、 $N_{\text{MDE}}$  の範囲が最も広い領域を採用した。なお、kNN 方向再構成法において  $\cos$  閾値を 0.95 とした条件では、上記の定義をそのまま適用すると全ての  $N_{\text{MDE}}$  が有効領域に含まれる場合がある。そのため、同一の定義基準を維持する目的で、当該条件に限りより厳しい 2.5%の

閾値を適用した。最後に、この過程から抽出される各 SN 距離毎の  $N_{\text{MDE}}$  領域の中央値を算出し、平均事象数と関連づける。高エネルギー事象および全エネルギー事象に対し、単純平均および kNN 方向再構成手法を用いて、以上の手順に従って抽出した  $N_{\text{MDE}}$  連続領域の中央値と、平均事象数との関係を図 5.23 と図 5.24 に示す。これらの分布から、抽出された  $N_{\text{MDE}}$  と事象数には完全ではないものの、べき乗則的な傾向が見られると判断した。そのため、他のパラメータ推定と同様に、式 5.4 のべき乗関数を用いてフィッティングを行い、 $N_{\text{MDE}}$  の推定式を構築した。高エネルギー事象および全エネルギー事象に対する各方向再構成手法ごとの  $N_{\text{MDE}}$  の推定式を表 5.7 にまとめる。

表 5.7: 高エネルギー事象に対する各フィッティングパラメータの推定式

方向再構成方法	対象エネルギー事象	$N_{\text{MDE}}$ の推定式
単純平均	高エネルギー	$0.168(N_{\text{events}})^{0.938}$
kNN(cos_th=0.85)	高エネルギー	$0.093(N_{\text{events}})^{1.006}$
kNN(cos_th=0.90)	高エネルギー	$0.063(N_{\text{events}})^{1.050}$
kNN(cos_th=0.95)	高エネルギー	$0.045(N_{\text{events}})^{1.094}$
単純平均	全エネルギー	$0.113(N_{\text{events}})^{0.983}$
kNN(cos_th=0.85)	全エネルギー	$0.130(N_{\text{events}})^{0.957}$
kNN(cos_th=0.90)	全エネルギー	$0.096(N_{\text{events}})^{1.000}$
kNN(cos_th=0.95)	全エネルギー	$0.047(N_{\text{events}})^{1.079}$

本章では、第 6 章の結果で取り扱う MC データや性能評価指標、DBSCAN-fitter における各アルゴリズムのパラメータ推定について述べた。次章では、本章で構築したパラメータ推定結果や評価指標を用いて、DBSCAN-fitter の方向再構成精度、安定性、および処理時間について定量的な評価を行い、既存手法との比較を通じて本手法の有効性を検証する。

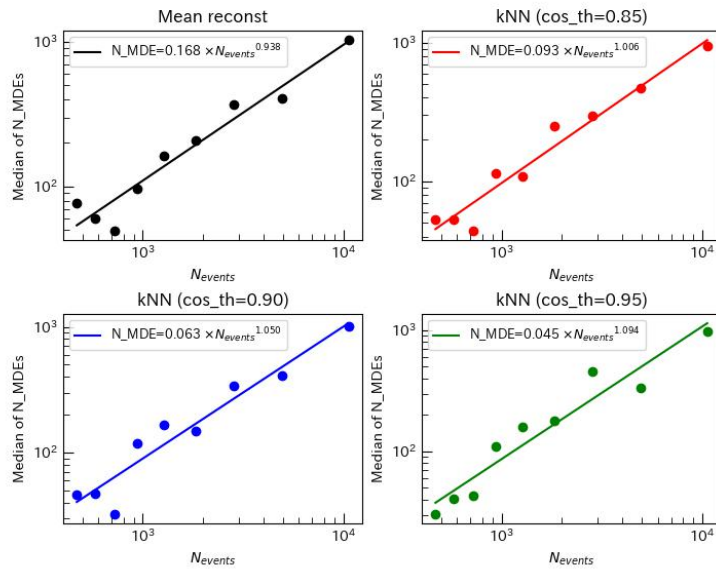


図 5.23: 高エネルギー事象において、安定な良好分解能を示す  $N\_MDE$  の連続領域の中央値と平均事象数の関係 (左上: 単純平均、右上:kNN (cos 閾値=0.85)、左下:kNN(cos 閾値=0.90)、右下:kNN(cos 閾値=0.95))

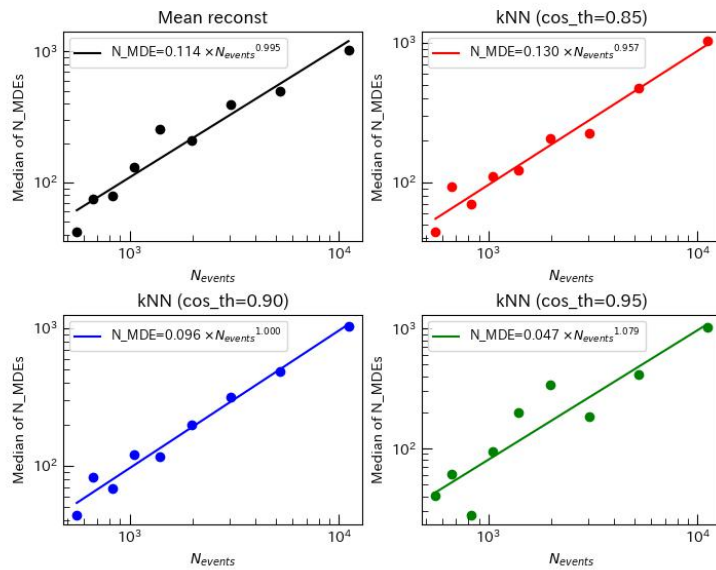


図 5.24: 全エネルギー事象において、安定な良好分解能を示す  $N\_MDE$  の連続領域の中央値と平均事象数の関係 (左上: 単純平均、右上:kNN (cos 閾値=0.85)、左下:kNN(cos 閾値=0.90)、右下:kNN(cos 閾値=0.95))

## 第6章 結果

本章では、第5章で構築した各種パラメータ推定式を用いて DBSCAN-fitter を適用し、その方向再構成性能を定量的に評価する。評価は角度分解能、Failure-rate、処理時間の三つの観点から行う。さらに、複数パラメータセットの導入による方向再構成の安定性向上についても検証する。さらに本研究では、これまでの解析で除外されてきた低エネルギー事象を解析に加える影響を調べることも目的の一つとしている。そのため、本章では、6 MeV より高いエネルギーを持つ事象のみを用いた解析結果 (以下、高エネルギー事象) と、6MeV 以下の事象を含めた解析結果 (以下、全エネルギー事象) の両方を示す。ここで「高エネルギー事象」および「全エネルギー事象」とは、それぞれ第3章で述べた事象選別を適用した後に残った事象を指す。

### 6.1 DBSCAN の性能評価

#### 6.1.1 複数パラメータセット導入による安定性の向上

本項では表 5.3 と表 5.4 の推定式を用いて、複数の DBSCAN パラメータセットを導入することによるクラスタ抽出の安定性向上の効果を評価する。

まず高エネルギー事象に対して、Middle\_eps(式 5.5) とそれに対応する min\_samples(式 5.2) のみを用いた場合の Not-finding-rate と、三種類の eps (Up\_eps, Middle\_eps, Bottom\_eps) とそれぞれに対応する min\_samples を用いた場合の Not-finding-rate の比較を行った。各 SN 距離毎に作成した 200 個のテスト用 MC データに DBSCAN を適用した際の Not-finding-rate を表 6.1 に示す。

Middle\_eps のみを用いた場合、SN 距離が 10kpc 以下における Not-finding はほぼ 0%である一方、18kpc 以上の遠距離では Not-finding-rate が 10%を超えることが分かる。これは、SN までの距離が増加するにつれて観測事象数が減少し、それに伴って ES 事象の局所密度が低下するため、単一のパラメータセットではクラスタ抽出に失敗する SN 事象が増加することに起因すると考えられる。この量は、銀河内で数十年から百年程度の頻度で発生するとされる SN ニュートリノバースト観測を想定した場合、十分に低いとは言い難い。一方、三種類の eps とそれぞれに対応する min\_samples を用いた場合には、18kpc 以上の遠距離においても、Not-finding-rate が大きく低減していることが確認できる。従って、複数のパラメータセットの導入は、特に遠距離発生した SN ニュートリノバースト事象に対し、クラスタ抽出の面において有効であると分かる。

表 6.1: 高エネルギー事象、200 個のテストデータに DBSCAN を適用した際の Not-finding-rate

SN 距離	Middle_eps とそれに対する min_samples による Not-finding-rate [%]	3 組のパラメータセットによる Not-finding-rate [%]
4kpc	0.0	0.0
6kpc	0.0	0.0
8kpc	0.0	0.0
10kpc	0.3	0.0
12kpc	3.3	0.5
14kpc	4.0	1.0
16kpc	7.7	1.3
18kpc	10.3	3.7
20kpc	14.3	4.7

次に、Not-finding-rate の更なる低減を目的として、各 eps に対応する min\_samples の推定値の周辺に追加の min\_samples を導入する手法について検討した。実際の SN ニュートリノバースト事象や個々の MC データにおいては ES 事象の分布構造や密度に統計的な揺らぎが存在する。DBSCAN は、min\_samples を閾値としてクラスタ形成の成否が決定される手法であるため、推定値近傍におけるわずかな条件差によっても、クラスタ抽出の成否が変化する可能性がある。そのため、単一の推定値に依存するのではなく、この推定値の周辺に位置する min\_samples を追加で用いることで、クラスタ抽出の成功率を維持または向上させつつ、方向再構成精度のよりパラメータセットが含まれる確率を高められると考えた。

この周辺値を定量的に評価するために、各 eps ごとに推定される min\_samples を基準値とし、その値を一定割合だけ増減させた二種類の min\_samples を追加し、合計 9 組のパラメータを用いて訓練データに対する解析を行った。具体的には、各 eps に対して推定される min\_samples を  $m$  としたとき、 $(1 - \alpha)m$  および  $(1 + \alpha)m$  ( $\alpha = 0.05 \sim 0.20$ ) を追加の min\_samples として用いた。この操作により三種類の eps とそれぞれに対応する min\_samples からなる合計 9 組のパラメータセットが構築される。高エネルギー事象に対して、 $\alpha$  を変化させながら 9 組のパラメータセットを用いた場合の Not-finding-rate を図 6.1 に示す。また、全エネルギー事象に対して同様の解析を行った結果を図 6.2 に示す。なお、本調査は表 6.1 において Not-finding-rate が 0% に至っていない場合があった遠距離 (10-20kpc) を対象として行った。高エネルギー事象および全エネルギー事象の両方とも、 $\alpha = 0.075$  のときに、いずれの距離においても安定して低い Not-finding-rate を得られることが確認できる。以上の結果を踏まえ、本研究では推定される min\_samples に対して  $\times 0.925$  および  $\times 1.075$  した二種類の min\_samples を追加し、合計 9 組の DBSCAN パラメータセットを用いて SN ニュートリノ事

象のクラスタ抽出を行う方針とした。

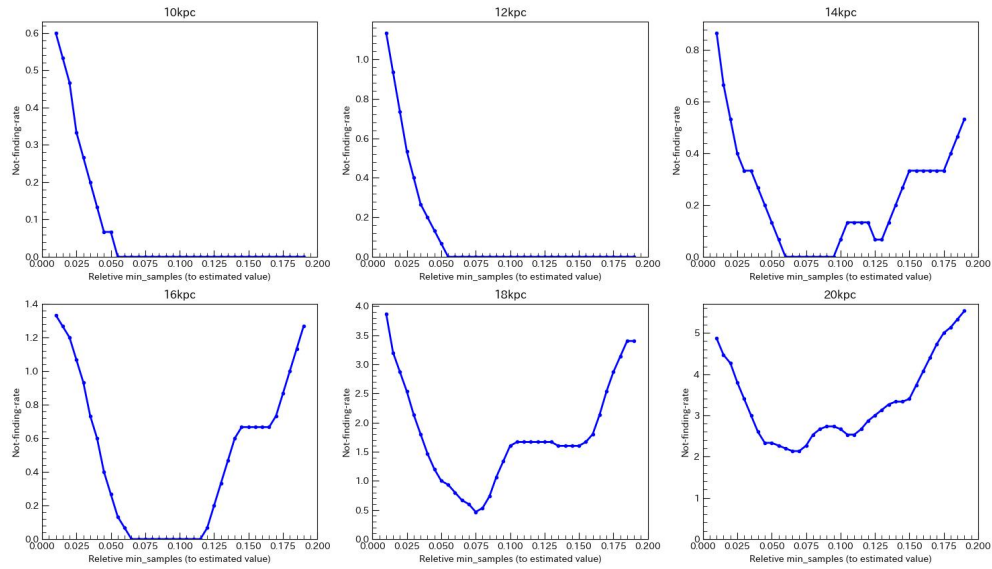


図 6.1: 高エネルギー事象において、推定される `min_samples` を基準とし、その値を一定割合だけ増減させた二種類の `min_samples` を追加した場合の Not-finding-rate の比較。横軸は追加した `min_samples` が推定値に対して何%に相当するかを示す)

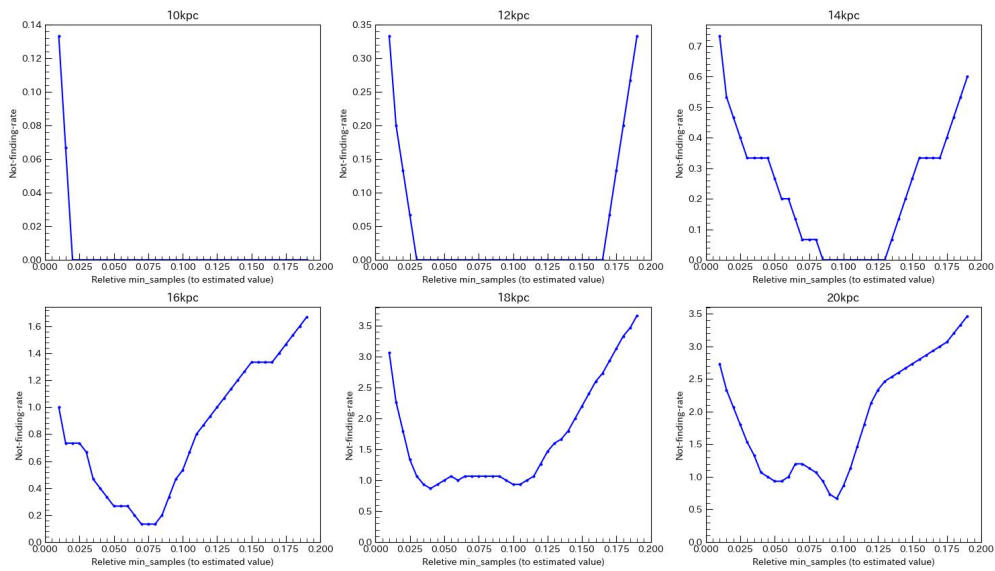


図 6.2: 全エネルギー事象において、推定される `min_samples` を基準とし、その値を一定割合だけ増減させた二種類の `min_samples` を追加した場合の Not-finding-rate の比較。横軸は追加した `min_samples` が推定値に対して何%に相当するかを示す

### 6.1.2 最終 Not-finding-rate

本項では、高エネルギー事象および全エネルギー事象に対して 9 組のパラメータセットを用いた場合の最終的な Not-finding-rate を示す。算出には、本研究で用いた全 MC データを用いた。その結果を表 6.2 および図 6.3 に示す。最終的に、SN 距離 20kpc における Not-finding-rate はいずれの場合も約 2.5%となった。この値は、データ数が異なる点に注意する必要があるものの、表 6.1 に示した 3 組のパラメータセットを用いた場合の 20kpc における値と比較して、約 2%の低減に相当する。以上より、パラメータセット数を増加させることで、Not-finding-rate をさらに低減し、クラスタ抽出の安定性を高めることができた。

本研究では、kNN 方向再構成と Estimator の計算コストとのバランスを考慮し、使用するパラメータセット数を 9 組とした。一方で、計算時間の増加を許容する場合には、パラメータセット数を 15 組程度まで拡張することで、クラスタ抽出の成功率をさらに向上できる可能性がある。

表 6.2: 高エネルギー事象および全エネルギー事象に対して 9 組のパラメータセットを用いた際の最終的な Not-finding-rate

SN 距離	総 MC データ数	高エネルギー事象に対する Not-finding-rate [%]	全エネルギー事象に対する Not-finding-rate [%]
4kpc	597	0.0	0.0
6kpc	789	0.0	0.0
8kpc	798	0.0	0.0
10kpc	919	0.0	0.0
12kpc	927	0.0	0.0
14kpc	947	0.1	0.1
16kpc	963	0.3	0.3
18kpc	978	0.7	0.9
20kpc	978	2.7	2.2

## 6.2 DBSCAN-fitter 後半処理の評価

本節では、DBSCAN によるクラスタ抽出後に適用される kNN 方向再構成法および Minimum Distance Estimator(MDE) の性能評価について述べる。また、kNN 方向再構成法については、

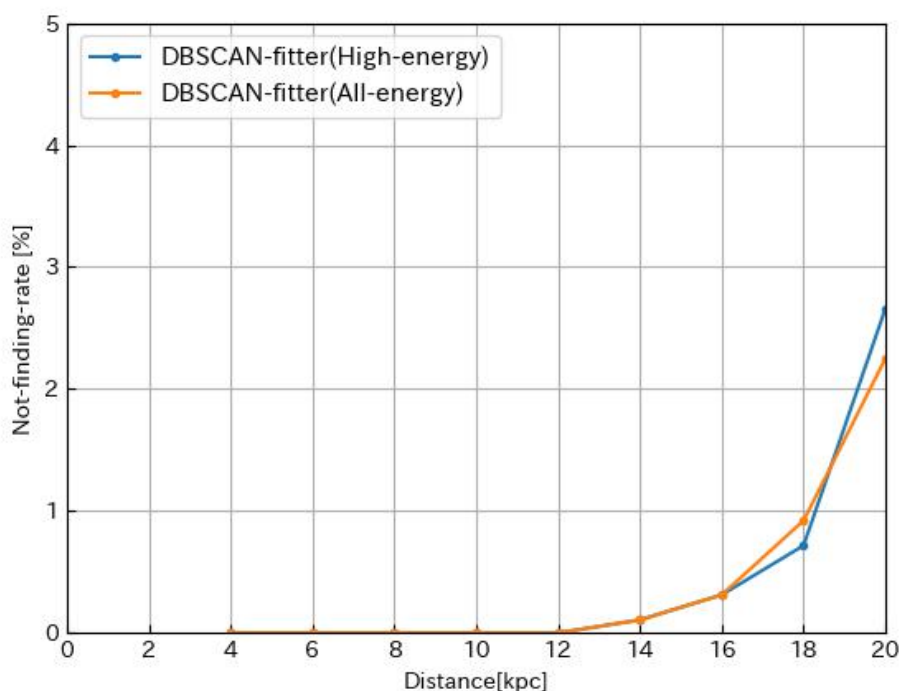


図 6.3: 高エネルギー事象および全エネルギー事象に対する最終的な Not-finding-rate。両者とも 20kpc におけるクラスタ抽出の失敗率が 2.5%程度となった。

有効性を検証するために、DBSCAN により抽出されたクラスタ内事象の運動方向を単純平均する手法による結果とも比較を行う。

### 6.2.1 kNN 再構成法における $\cos$ 閾値による角度分解能の変化

高エネルギー事象および全エネルギー事象について、全 MC データに対して DBSCAN-fitter を適用し、5.4 節で設定した kNN 再構成方法における各  $\cos$  閾値の角度分解能を算出した。この結果から、kNN 方向再構成手法の有効性および最適な  $\cos$  閾値についての検証を行った。得られた角度分解能を表 6.3、表 6.4、図 6.4、図 6.5 に示す。

これらの結果から、高エネルギー事象および全エネルギー事象のいずれにおいても、DBSCAN によって抽出されたクラスタ内事象の運動方向を単純平均する手法と比較して、kNN 再構成手法の方が良好な角度分解能を示すことが分かる。特に、18kpc 以上の遠距離では、角度分解能が  $0.3^\circ$  以上改善されており、kNN 再構成法が遠距離における再構成性能の向上に大きく寄与していることが分かる。次に、 $\cos$  閾値の違いによる角度分解能への影響に着目すると、いずれのエネルギー事象においても、 $\cos$  閾値を 0.95 とした場合に最良な角度分解能が得られた。これは、過度に広い角度範囲を許容することによるバックグラウンド事象の混

入を抑制できたことに加え、局所密度が高く SN ニュートリノ到来方向と強く相関する事象が方向再構成に適切に反映されたためと考えられる。以上の結果を踏まえ、本研究における最終的な角度分解能の評価には、 $\cos$  閾値 0.95 の kNN 方向再構成手法による結果を採用する。

表 6.3: 高エネルギー事象における単純平均手法と kNN 再構成手法角度分解能の比較

SN 距離	角度分解能	角度分解能	角度分解能	角度分解能
	$1\sigma$ [deg] 単純平均	$1\sigma$ [deg] kNN 再構成法 ( $\cos\_th = 0.85$ )	$1\sigma$ [deg] kNN 再構成法 ( $\cos\_th = 0.90$ )	$1\sigma$ [deg] kNN 再構成法 ( $\cos\_th = 0.95$ )
4kpc	1.90±0.05	1.92±0.05	1.90±0.05	1.84±0.05
6kpc	2.74±0.05	2.65±0.07	2.66±0.05	2.55±0.06
8kpc	3.65±0.10	3.65±0.08	3.64±0.09	3.53±0.12
10kpc	4.54±0.13	4.45±0.13	4.40±0.11	4.31±0.11
12kpc	5.43±0.14	5.32±0.12	5.27±0.13	5.26±0.13
14kpc	6.11±0.15	6.00±0.17	5.87±0.12	5.81±0.15
16kpc	7.35±0.18	7.31±0.14	7.18±0.15	7.01±0.12
18kpc	9.20±0.28	8.77±0.26	8.63±0.30	8.64±0.29
20kpc	10.65±0.30	10.37±0.25	10.25±0.29	10.07±0.31

表 6.4: 全エネルギー事象における単純平均手法と kNN 再構成手法角度分解能の比較

SN 距離	角度分解能	角度分解能	角度分解能	角度分解能
	$1\sigma$ [deg] 単純平均	$1\sigma$ [deg] kNN 再構成法 ( $\cos\_th = 0.85$ )	$1\sigma$ [deg] kNN 再構成法 ( $\cos\_th = 0.90$ )	$1\sigma$ [deg] kNN 再構成法 ( $\cos\_th = 0.95$ )
4kpc	1.92±0.05	1.96±0.05	1.90±0.05	1.89±0.06
6kpc	2.70±0.07	2.76±0.06	2.68±0.05	2.66±0.05
8kpc	3.78±0.11	3.76±0.08	3.68±0.10	3.57±0.08
10kpc	4.45±0.11	4.46±0.11	4.36±0.09	4.34±0.10
12kpc	5.52±0.11	5.33±0.12	5.25±0.11	5.23±0.11
14kpc	6.28±0.15	6.19±0.14	6.18±0.16	5.94±0.13
16kpc	7.70±0.12	7.47±0.15	7.32±0.16	7.34±0.16
18kpc	9.42±0.20	9.38±0.26	9.22±0.22	9.07±0.20
20kpc	11.13±0.27	10.95±0.35	10.77±0.30	10.59±0.32

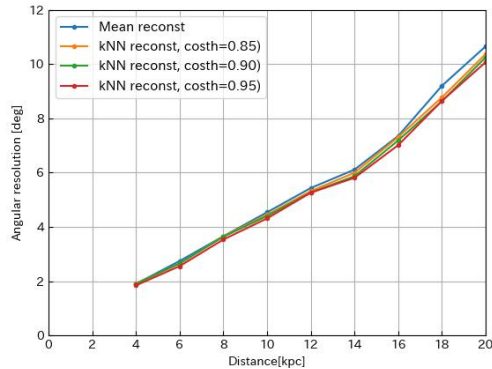


図 6.4: 高エネルギー事象に対して DBSCAN-fitter を適用した場合における各  $\cos$  閾値の角度分解能。単純平均による再構成法 (青線) よりも、 $\cos$  閾値=0.95 での kNN 再構成法 (赤線) の方がより良い分解能を示す。

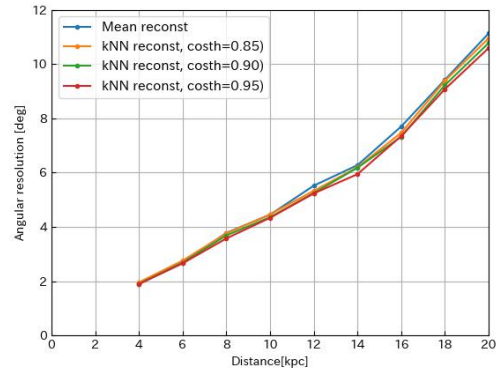


図 6.5: 全エネルギー事象に対して DBSCAN-fitter を適用した場合における各  $\cos$  閾値の角度分解能。単純平均による再構成法 (青線) よりも、 $\cos$  閾値=0.95 での kNN 再構成法 (赤線) の方がより良い分解能を示す。

## 6.2.2 Estimator の性能評価

高エネルギー事象および全エネルギー事象を用いて、本研究で導入した Minimum Distance Estimator(MDE) の性能評価を行った。本評価では、MDE が 9 通りの再構成結果の中から、明らかに不適切な結果を回避できるかに着目し、その失敗割合を Estimator Failure-rate として定義した。Estimator Failure-rate の算出においては、各 MC データから得られる 9 通りの再構成結果のうち、真の SN 方向と再構成方向との角度差  $\Delta\theta_{SN}$  の最悪値が表 5.1 に示す閾値を超える SN 事象のみを対象とした。これは、最悪値が閾値以下である場合には、Estimator の有無に関わらず十分な再構成性能が得られていると判断できるためである。評価対象となる各 MC データについて、角度差  $\Delta\theta_{SN}$  の最良値と最悪値の中間値を基準値と定義し、MDE によって選択された再構成結果の角度差  $\Delta\theta_{SN,MDE}$  がこの基準値を上回った場合を「失敗」と判定した。この判定方法に基づいて失敗数をカウントし、各 SN 距離における Estimator Failure-rate を、全 MC データに対する失敗数の割合として算出した。その結果を図 6.6 に示す。なお、MDE の性能評価には、各パラメータセット間で角度差  $\Delta\theta_{SN}$  の差が顕著に現れる単純平均による方向再構成結果を用いた。また、評価にはパラメータ最適化に用いた訓練データとは独立データセットとして、各 SN 距離につき 200 個の MC データをテスト用に使用した。図 6.6 から、Estimator Failure-rate は約 20-40%であり、約 60%以上の SN 事象に対して、MDE が明らかに不適切な再構成結果を回避できていることが分かった。また、SN 距離が大きくなるにつれて、Estimator Failure-rate が低下する傾向が見られた。遠距離 SN では事象数の減少に伴い ES 事象が集中する領域の局所密度が低下するため、適切なクラスタを与えるパ

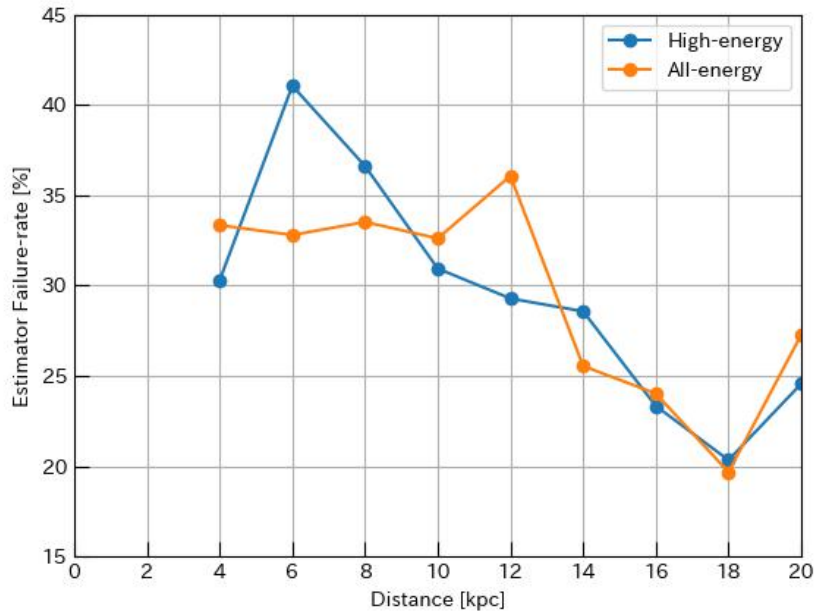


図 6.6: 高エネルギー事象と全エネルギー事象における Estimator Failure-rate。青線が高エネルギー事象による結果、オレンジ線が全エネルギー事象による結果を示す

ラメータセットが限定される。その結果、9 組の再構成結果の中に不適切な再構成結果が含まれる割合が高まる。にもかかわらず Estimator Failure-rate が低下していることから、MDE は不適切な再構成結果を回避する上で有効に機能していると結論づけられる。

次に、MDE が最良の再構成結果をどの程度選択できているかを確認するために、各 MC データ毎の最良の角度差  $\Delta\theta_{SN,best}$  と、MDE によって選択された結果の角度差  $\Delta\theta_{SN,MDE}$  との関係を示す散布図を作成した。その結果を図 6.7 に示す。

MDE が常に最良の再構成結果を選択している場合、両者の角度差は一致し、データ点は  $y = x$  の直線上に分布するはずである。しかし、図 6.7 が示すように、多くのデータ点がこの直線の上側に分布していることから、MDE は大部分の SN 事象に対して最良の再構成結果を選択できていないことが分かる。実際、全 SN 事象のうち最良の再構成結果を選択できた割合は、高エネルギー事象および全エネルギー事象のいずれにおいても 21.2%にとどまった。

以上の結果をまとめると、MDE は明らかに不適切な再構成結果を回避する能力を有する一方で、最良の再構成結果を高い確率で選択するには至っていないことが分かった。MDE は再構成方向周辺の局所的な事象密度に基づいて評価を行うため、誤った方向に再構成された場合には、局所密度の低下を通じてその状況を適切に評価できる。一方で、MDE は再構成方向近傍における事象分布の形状や広がりといった分布特性を考慮できない。そのため、真の SN 方向から離れた位置に ES 事象が局所的に集中した場合であっても、その分布の偏りを適切に評

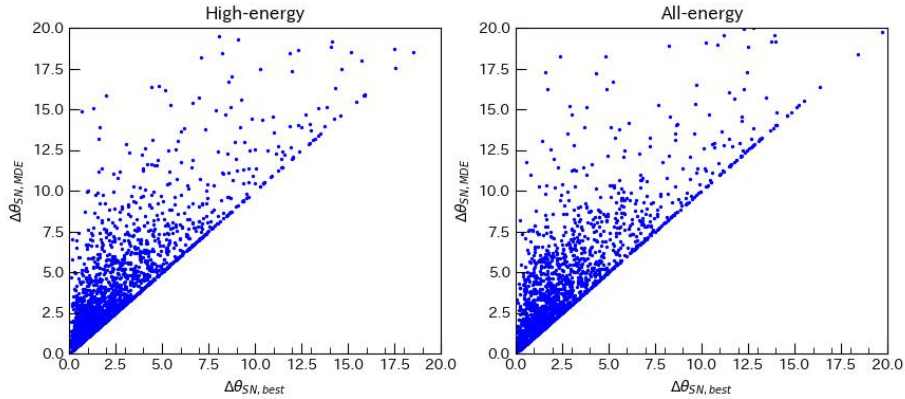


図 6.7: 最良の角度差  $\Delta\theta_{SN,best}$  と MDE によって選択された結果の角度差  $\Delta\theta_{SN,MDE}$  との関係を示す散布図 (左: 高エネルギー事象、右: 全エネルギー事象)。MDE は多くの SN 事象に対して最良の角度差を持った結果を抽出できていないことを示している。

価できず、結果として角度差  $\Delta\theta_{SN}$  がより小さい再構成結果を選択できない場合が生じる。以上の MDE の特性が、最良の再構成結果の選択率が低い要因であると考えられ、今後の改良に向けた重要な課題である。

### 6.3 角度分解能の比較

全 MC データを用いて、高エネルギー事象・全エネルギー事象に対する事象選別を行い、これまでに決定した各種パラメータを用いて DBSCAN-fitter を適用した場合の角度分解能を算出した。さらに、得られた角度分解能について、現行の SNWATCH で採用されている SN 方向再構成手法との比較を行った。これらの結果を図 6.8 および表 6.5 に示す。

図 6.8 から、DBSCAN-fitter による角度分解能は、SN 距離が大きくなるにつれて悪化する傾向が確認された。これは、距離の増加に伴って全事象数が減少することにより、SN ニュートリノ到来方向と強く相関する ES 事象の数も低下したからだと考えられる。

次に、図 6.8 と表 6.5 より、DBSCAN-fitter の角度分解能は、現行の SNWATCH で採用されている方向再構成手法と比較して全体的に劣っていることが分かった。例えば、SN 距離 10kpc における DBSCAN-fitter は約  $4.3^\circ$  であるのに対し、現行手法は  $3.75^\circ$  であり、約  $0.6^\circ$  の差がある。同様に、SN 距離 4kpc では約  $0.4^\circ$ 、20kpc では約  $1.5^\circ$ 、DBSCAN-fitter の分解能が悪化している。

DBSCAN-fitter は、ES 事象が集中したクラスタ領域を直接抽出できる点で、現行の SN 方向再構成手法とは異なる特性を有する。一方で、限定的な領域のみを抽出することにより、抽出された事象群の分布が SN の真の方向に対して偏る場合がある。特に、事象の集中度が低下する遠距離では、このような偏りが生じやすい。本研究では、kNN 方向再構成手法を用いるこ

表 6.5: DBSCAN-fitter と現行の SNWATCH 方向再構成法による最終的な角度分解能の比較

SN 距離	角度分解能 $1\sigma$ [deg]	角度分解能 $1\sigma$ [deg]	角度分解能 $1\sigma$ [deg]
	SNWATCH[17]	DBSCAN-fitter	DBSCAN-fitter
	IBD 事象タグ付け率:50%	IBD 事象タグ付け率:55%	IBD 事象タグ付け率:55%
	7MeV threshold 2m Dwall cut	高エネルギー事象選別 kNN(cos_th = 0.95)	全エネルギー事象選別 kNN(cos_th = 0.95)
4kpc	1.48	1.84±0.05	1.89±0.06
6kpc	2.26	2.55±0.06	2.66±0.05
8kpc	3.05	3.53±0.12	3.57±0.12
10kpc	3.75	4.31±0.11	4.34±0.10
12kpc	4.66	5.26±0.13	5.23±0.11
14kpc	5.57	5.81±0.15	5.94±0.13
16kpc	6.49	7.01±0.14	7.35±0.16
18kpc	7.74	8.64±0.29	9.07±0.20
20kpc	8.85	10.07±0.31	10.59±0.32

とで、より局所的に集中度の高い事象を重く評価する工夫を行った。しかし、この方法のみでは抽出事象分布に残存する偏りを完全に補正することは困難であるため、SNWATCH と比較して方向再構成精度が低下したと考えられる。また、6.2.2 項で述べたように、最終結果の選別に用いる MDE も抽出事象分布の偏りを評価する性能を有していない。したがって、良好な方向再構成を抽出できず、角度分解能が SNWATCH に比べて劣る結果になった可能性がある。

次に、DBSCAN-fitter における高エネルギー事象のみを用いた場合と、そこに低エネルギー事象を含めた場合の分解能を比較すると、4-12kpc の距離範囲では両者に大きな差は見られない一方で、18kpc 以上の遠距離においては分解能に  $0.4^\circ$  以上の差が生じている。遠距離 SN の低エネルギー領域では、ES 事象数が元々少ないため、各種事象情報による選別後でも ES 事象の割合は近距離 SN と比べて低くなる。実際、4kpc と 20kpc の MC データについて、低エネルギー事象に対する選別後の ES 事象割合を調べたところ、4kpc では平均 21%であったのに対し、20kpc では 5%であった。したがって遠距離 SN では、方向推定におけるバックグラウンド事象の混入が増加することで、角度分解能が悪化したと考えられる。

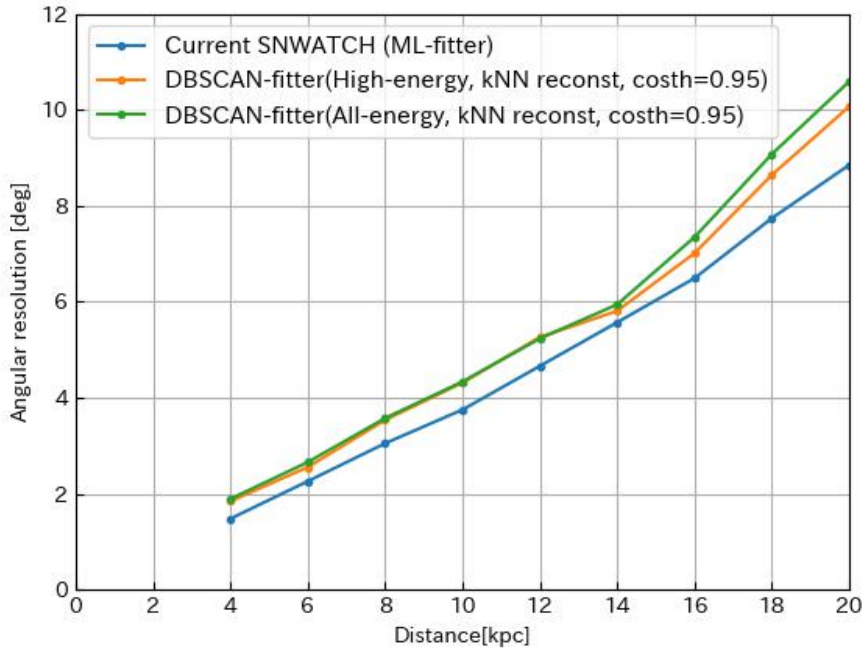


図 6.8: DBSCAN-fitter と現行の SNWATCH 方向再構成法による最終的な角度分解能の比較

## 6.4 Failure rate の比較

全 MC データを用いて、高エネルギー事象・全エネルギー事象に対する事象選別を行い、これまでに決定した各種パラメータを用いて DBSCAN-fitter を適用した場合の Failure-rate を算出した。Failure-rate は 5.2.2 項で述べた手順に従って算出される。その結果を表 6.6 および図 6.9 に示す。

DBSCAN-fitter による Failure-rate は、Not-finding-rate と同様に、SN 距離が 4-12kpc までは 0.0%を示し、14kpc 以降では距離の増加と伴って徐々に増加する傾向が見られた。DBSCAN-fitter と SNWATCH における現行手法の Failure-rate を比較すると、14kpc までは両者はほぼ同等である一方で、16kpc 以上の遠距離では DBSCAN-fitter の方が全体として低い Failure-rate を示すことが分かる。これは、複数の DBSCAN パラメータセットの導入にすることで、遠距離 SN においてもクラスタ抽出に適したパラメータが含まれる可能性が高まったことに加え、6.2.2 項で述べたように MDE が明らかに不適切な再構成結果を回避するため、低い Failure-rate が達成されたと考えられる。

表 6.6: DBSCAN-fitter と現行の SNWATCH 方向再構成法による最終的な Failure-rate の比較

SN 距離	Failure-rate [%]	Failure-rate [%]	Failure-rate [%]
	SNWATCH[17]	DBSCAN-fitter	DBSCAN-fitter
	7MeV threshold	高エネルギー事象選別	全エネルギー事象選別
	2m Dwall cut	kNN(cos_th = 0.95)	kNN(cos_th = 0.95)
4kpc	0.0	0.0	0.0
6kpc	0.0	0.0	0.0
8kpc	0.0	0.0	0.0
10kpc	0.0	0.0	0.0
12kpc	0.1	0.0	0.0
14kpc	0.2	0.2	0.2
16kpc	1.1	0.6	0.6
18kpc	3.5	1.8	2.5
20kpc	5.5	5.7	4.3

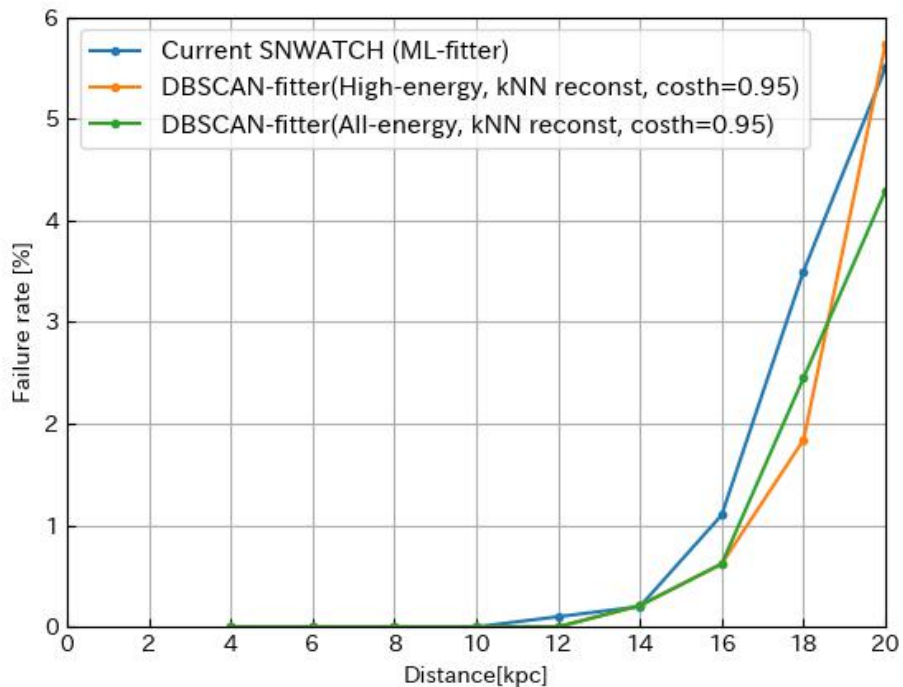


図 6.9: DBSCAN-fitter と現行の SNWATCH 方向再構成法による最終的な Failure-rate の比較

## 6.5 Reconstruction time

全 MC データを用いて、全エネルギー事象に対する事象選別を行い、これまでに決定した各種パラメータを用いて DBSCAN-fitter を適用した場合の処理時間を算出した。その結果を表 6.7 に示す。なお、ここで示す処理時間には、WIT による事象情報の再構成時間は含まれていない。

表 6.7: 全エネルギー事象に対する DBSCAN-fitter の処理時間

SN 距離	平均事象数	平均処理時間 [sec]
4kpc	11109	1.26
6kpc	5222	0.42
8kpc	3022	0.19
10kpc	1981	0.11
12kpc	1395	0.06
14kpc	1050	0.05
16kpc	826	0.04
18kpc	670	0.03
20kpc	558	0.02

SNWATCH における現行手法では、事象数が 3000 以下の SN 事象に対して、解析は 1.0 秒以内に終了する [17]。一方、DBSCAN-fitter では、同程度の事象数を持つ 8kpc の場合において、処理時間は約 0.2 秒であり、現行手法と比較して約 5 倍高速に解析を行うことができる。また、事象数が約 10000 の場合、現行の手法では解析に約 2.2 秒 [17] を要するのに対し、DBSCAN-fitter では、それよりも多い事象数を含む 4kpc において約 1.3 秒で解析が完了する。この結果から、DBSCAN-fitter では、事象数の増加に対しても処理時間が比較的穏やかに増加することが分かる。これは、DBSCAN とその後段の方向再構成処理が局所的な距離計算や単純なベクトル演算を中心とした比較的軽量な計算で構成されていることに起因すると考えられる。

以上より、DBSCAN-fitter は、高速な方向再構成が要求される超新星アラートシステムにおいて、計算時間の観点からも実運用に適した手法であると結論づけられる。

## 第7章 まとめ・今後の展望

本研究では、スーパーカミオカンデにおける超新星ニュートリノバースト観測を対象として、実際の超新星爆発時に再構成される事象をモンテカルロシミュレーションによって再現し、これまでの SNWATCH では除去されてきた 6MeV 以下の事象も含めた解析を行った。その上で、DBSCAN に基づくクラスタ抽出と後段の方向再構成手法を統合した解析システムである DBSCAN-fitter を構築し、その性能評価を行った。

各種パラメータの最適化を行った結果、遠距離に位置する超新星に対しても高い確率で電子散乱 (ES) 事象由来のクラスタを抽出できることが確認された。一方で、最終的に得られる角度分解能は、現行の SNWATCH における方向再構成手法と比較して同等の性能には至らなかった。一方、Failure-rate に関しては、特に 16kpc 以上の遠距離において現行手法よりも低い値を示すことが確認された。また、方向再構成に要する処理時間についても、現行手法と比較して高速に解析が可能であり、DBSCAN-fitter が持つ計算効率の高さが示された。これらの点から、本手法は角度分解能の観点では課題が残るものの、安定性および即時性の面で有意な利点を有することが明らかとなった。

今後の展望としては、DBSCAN-fitter における角度分解能のさらなる向上が挙げられる。本研究では、DBSCAN パラメータの最適化や kNN による局所密度を考慮した方向再構成を導入することで、10kpc における角度分解能を約  $4.3^\circ$  付近まで改善することができたが、特に最終的な結果選択を担う Minimum Distance Estimator(MDE) には依然として改良の余地が残されている。今後は、再構成方向周辺の局所密度に加え、クラスタ内の分布形状を考慮できる指標の導入や、パラメータセット数の増加による探索範囲の増加を通じて、さらなる性能の向上が期待される。

# 付録 A SNWATCH における最大尤度 (ML) 方向再構成アルゴリズム

SNWATCH では、超新星ニュートリノバーストに対して最大尤度法 (Maximum Likelihood:ML) を用いた方向再構成を行う [11]。ML-fitter では、各事象に対して反応チャンネルの混合モデルとして尤度  $L_i$  を次式で定義する。

$$L_i = \sum_r N_{r,k} t_r(f_i) p_r(E_i, \hat{d}_i; \hat{d}_{SN}^{recon}) \quad (7.1)$$

表 7.1: ML-fitter における SN ニュートリノ反応チャンネル

r-bin	反応チャンネル	k-bin	エネルギー範囲 [MeV]
0	IBD	0	$6 < E < 10$
1	(反電子ニュートリノによる)ES	1	$10 < E < 15$
2	(他ニュートリノによる)ES	2	$15 < E < 22$
3	O16	3	$22 < E < 35$
		4	$35 < E < 50$

ここで、 $r$  は反応チャンネル、 $k$  は事象  $i$  の再構成エネルギー  $E_i$  により決まるエネルギービンを表す (表 7.1)。また、 $E_i$  は事象の再構成エネルギー、 $\hat{d}_i$  は各事象の再構成運動方向、 $\hat{d}_{SN}^{recon}$  は求める SN 方向である。 $p_r(E_i, \hat{d}_i; \hat{d}_{SN}^{recon})$  は確率密度関数 (PDF) であり、主に  $\hat{d}_i \cdot \hat{d}_{SN}^{recon} = \cos \theta_i$  に依存し、各反応チャンネルに対する角度分布 (前方ピーク/等方成分) を反映する。PDF は MC データに基づいて決定され、各エネルギービンにおける角度分布をモデル関数でフィットすることで与える。

現行の SNWATCH では、IBD 判定ラベル  $f_i$  を用いた重み  $t_r(f_i)$  を尤度に導入している。IBD チャンネルに対しては

$$t_{r=IBD}(f_i) = \begin{cases} \epsilon_{tag} & (f_i = \text{TRUE}) \\ 1 - \epsilon_{tag} & (f_i = \text{FALSE}) \end{cases} \quad (7.2)$$

と定義し、IBD タグ付け効率  $\epsilon_{tag}$  を用いて IBD 由来確率を尤度に反映させる。これにより、方向情報をほとんど持たない IBD 事象の寄与を適切に扱い、ES 事象をより有効に抽出できる。

バースト全体の尤度は

$$L = \exp(\sum_{r,k} N_{r,k}) \prod_i L_i \quad (7.3)$$

と定義され、 $L$  を最大化することで再構成方向  $\hat{d}_{SN}^{recon}$  を決定する。

## 謝辞

本研究を進めるにあたり、多くの方々に多大なるご指導とご協力を賜りました。その皆様にこの場を借りて深く感謝申し上げます。指導教員の南野教授には、お忙しい中にも関わらず、研究相談をはじめ、学会発表や論文作成に至るまで多くのお時間を割いて、ご指導いただきました。時には無理なお願いをしてご迷惑をおかけすることもありましたが、いつも親身に対応してくださいました。この3年間、本研究に楽しく励むことができたのは、ひとえに南野教授のおかげであり、心より感謝申し上げます。

また、本研究において深くご指導いただいた Lluis 先生には、研究で行き詰まった際や不明な点が生じた時に、常に丁寧に対応していただきました。いつも拙い英語で申し訳ないと感じておりましたが、辛抱強く議論を重ねてくださり、理解を深めようとする姿勢から多くを学びました。本研究は、私が新しい技術に興味を抱いていたことと、Lluis さんが取り組みたいと考えていた内容が偶然にも一致したことを契機として始まりました。このような貴重な学びの機会を与えてくださったことに、心より感謝申し上げます。

池田先生、Guillaume 先生、関谷先生をはじめとする SNWATCH グループの皆様にも感謝申し上げます。本研究では複雑な内容を扱うことも多く、報告の際には分かりづらい点もあったかと思いますが、それにもかかわらず的確で有益な助言をいただきました。短い期間ではありましたが、心より感謝申し上げます。

同じ研究室のメンバーにも感謝申し上げます。就職された先輩の伊藤さんには、私が学部4年生の頃から親しくしていただき、研究の話だけでなく日常の雑談や休日の交流を通して、気分転換に付き合ってくださいました。佐々木くん、鈴木くん、高野さんは同期として研究に加えお互いの進路について話すこともあり、互いに刺激を受けながら研究生活を送ることができました。平田くんには研究だけでなく雑談でも非常に面白い話を聞かせてもらいました。秋の物理学会での発表後に一緒にお好み焼きを食べた一日は、特に印象深い思い出として残っています。猪俣くんと城下くんには、研究室内で様々なイベントを企画していただき、楽しい研究室生活を送ることができました。加藤さん、岩橋くん、北島くん、佐藤くん、橋戸くんとは、研究室行事などを通じて交流したり、時には趣味の話で盛り上がりたりもしました。中西さんには本研究の関連する研究を行っていただいたこともあり、同じ研究室で本研究について深く議論できたことに、心より感謝しています。中西さんの探究心や発想力にはいつも驚かされ、時には私が思いつきもしなかった興味深い意見を示してくださることもあり、大きな刺激を受けました。本研究が今後どのように発展していくのか非常に楽しみにしております。どうかお身体に気をつけて、研究に励んでほしいと思います。

最後に、家族には食事や家事など生活面だけでなく、精神面においても支えてもらいました。慌ただしい日々の中で変わらずに支援してくれたことに、心より感謝しています。

ここに記載した方々以外にも、多くの方々の支えがあって本研究を進めることができました。ご協力くださった皆様に、改めて深く御礼申し上げます。

## 参考文献

- [1] 山田 章一. **新天文学ライブラリー 4 超新星**. 日本評論社, 2016.
- [2] A. Burrows and D. Vartanyan. Core-collapse supernova explosion theory. *Nature*, 589(7840):29–39, January 2021.
- [3] Hans-Thomas Janka. Core-collapse supernovae: Reflections and directions. *Prog. Theor. Exp. Phys.*, 2012:01A309, 2012.
- [4] Y. Fukuda et al. The Super-Kamiokande detector. *Nucl. Instrum. Meth. A*, 501:418–462, 2003.
- [5] 東京大学宇宙線研究所. Super-kamiokande. [https://www-sk.icrr.u-tokyo.ac.jp/~hayato\\_s/super-k.html](https://www-sk.icrr.u-tokyo.ac.jp/~hayato_s/super-k.html).
- [6] 浜松ホトニクス株式会社. ニュートリノを探る. <https://www.hamamatsu.com/jp/ja/why-hamamatsu/exploring-neutrinos.html>.
- [7] 浜松ホトニクス株式会社. 光電子増倍管その基礎と応用. [https://www.hamamatsu.com/content/dam/hamamatsu-photonics/sites/documents/99\\_SALES\\_LIBRARY/etd/PMT\\_handbook\\_v4J.pdf](https://www.hamamatsu.com/content/dam/hamamatsu-photonics/sites/documents/99_SALES_LIBRARY/etd/PMT_handbook_v4J.pdf), 2017.
- [8] Giada Carminati. The new Wide-band Solar Neutrino Trigger for Super-Kamiokande. *Phys. Procedia*, 61:666–672, 2015.
- [9] John F. Beacom and Mark R. Vagins. Antineutrino spectroscopy with large water Čerenkov detectors. *Physical Review Letters*, 93(17), October 2004.
- [10] K. Abe et al. Second gadolinium loading to super-kamiokande. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1065:169480, August 2024.
- [11] K. Abe et al. Real-time supernova neutrino burst monitor at super-kamiokande. *Astroparticle Physics*, 81:39–48, August 2016.
- [12] L. N. Machado et al. Pre-supernova alert system for super-kamiokande. *The Astrophysical Journal*, 935(1):40, August 2022.
- [13] Michael Smy. Low Energy Event Reconstruction and Selection in Super-Kamiokande-III. In *30th International Cosmic Ray Conference*, volume 5, pages 1279–1282, 7 2007.
- [14] Y. Kashiwagi et al. Performance of sk-gd’s upgraded real-time supernova monitoring system, 2024.
- [15] M. Ikeda et al. Search for supernova neutrino bursts at super-kamiokande. *The Astrophysical Journal*, 669(1):519–524, November 2007.
- [16] A. C. Müller and S. Guido. *Python* **ではじめる機械学習** — *scikit-learn* で学ぶ特徴量エン

エンジニアリングと機械学習の基礎 一. オライリー・ジャパン, 2017.

[17] Super-Kamiokande Collaboration. Private communication, paper in preparation.